Strain relaxation via dislocation formation in strained semiconductor structures

Gosling, T. J.

Award date:
1994

Awarding institution:
University of Bath

Strain Relaxation via Dislocation Formation in Strained Semiconductor Structures

submitted by

T. J. Gosling

for the degree of PhD

of the

University of Bath

1994

Signature of Author .....T. J. Gosling.....

T. J. Gosling

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Summary

Semiconductor structures containing intentionally strained regions may be fabricated by advanced crystal growth techniques that allow perfect atomic registry to be maintained across interfaces between semiconductor materials that have different lattice sizes. The lattice mismatch between the materials results in regions of the structure being strained. Such structures are important because of the ability to control electronic properties of devices by the introduction of strain. The strain can also give rise to the formation of defects, and in particular dislocations. This thesis forms a study of the conditions under which dislocations are introduced in strained structures, and the extent to which they propagate and relax the strain. To aid this task three new solutions are obtained here:

(i) Closed-form formulae for the stresses and energy of arrays of misfit dislocations at the interface between a strained layer and substrate that are elastically anisotropic and have the same elastic constants. First-order correction terms are also obtained, to allow for a difference between the elastic constants of the layer and substrate.

(ii) A line-integral representation for the stresses due to an arbitrary dislocation in an isotropic half-space, with closed-form formulae for the stresses due to a dislocation half-line.

(iii) Closed-form formulae for the stresses due to a periodic array of mismatched inclusions in an isotropic elastic body.

Solution (i) is used to compare two approaches to the study of the propagation and interaction of dislocations in a strained layer. One approach considers the driving force on a single dislocation, tending to extend it, and the other approach considers conditions under which the total energy is minimized. The two approaches are used to make predictions about combinations of system parameters at which the onset of strain relaxation should occur and about the long-time limit of the strain relaxation process. These predictions are compared with experiment. The nucleation of new dislocations is then modelled with the aid of solutions (ii) and (iii). Solution (ii) is used to obtain values for the activation energies for buried loops and surface half loops, taking the effects of the free surface at the top of the layer rigorously into account. The results are used to explain the experimentally observed change in the dislocation distribution as the lattice mismatch is increased. Solution (iii) is used to model the nucleation of dislocations at non-planar features of the interface between the strained-layer and its substrate. A brief indication is also given of the importance of this solution for estimation of the electronic properties of buried strained quantum wires.

Contents

Acknowledgements

Thanks are due to many people, of whom only a selection are named now. Particular thanks must go to Suresh Jain, who enticed me into science in the first place, and whose vigorous style has always been an inspiration. Ron Bullough has also been instrumental in my development, by giving me the opportunity to spend summer vacations working with him and Suresh Jain at Harwell Laboratory, IMEC, and T-U Delft, and allowing me to benefit from his great experience. Thanks are also due him for introducing me to my supervisor John Willis, whose enthusiasm and ability inject excitement into research, and whose toleration of my more stupid questions and suggestions is remarkable.

Many people have helped to make my stay in Bath an exceptionally happy one, in particular my colleagues in 1W3.6, and my cohabitants in Oxford House and 2 Park Lane. Of these special mention is deserved by Sally for introducing me to Lynne, and by Lynne for putting up with me.

Finally I should like to thank my parents for their constant support, and the staff of the School of Mathematical Sciences for all of their help.

Publications

Chapter 2 is based on reference [27] and Chapter 4 is based on reference [28]. Many of the ideas in Chapter 3 were partially developed in references [26] and [25]. Finally, much of the work appearing in Chapter 5 appears in references [28] and [24].

Chapter 1

Introduction

The advance of epitaxial growth techniques for semiconductors has allowed the manufacture of high quality strained layers, which are now an integral part of many high-performance device structures. The simplest strained-layer structure comprises a thin semiconductor layer grown on a thick substrate of a semiconductor that has a different lattice constant. Important examples of such so-called heterostructures involve layers of InGaAs grown on GaAs or of GeSi on Si. In spite of the mismatch between the natural lattice sizes of the layer and substrate, advanced growth techniques enable the lattices to be in near-perfect atomic registry across the substrate-layer interface, provided the layer thickness is smaller than a mismatch-dependent critical thickness, which will be defined formally later. If the substrate is very much thicker than the layer and the lateral dimensions of the layer are much greater than its thickness, as is usually the case in device applications, then the lattice mismatch is accomodated entirely through strain in the layer, with the substrate remaining, to all intents and purposes, unstressed.

The strain in the layer affects the band structure there, and therefore variation of the strain, which is easily controlled by changing the composition and hence natural lattice constant of the layer material, opens up the possibility of tailoring the band gap of the layer, and the band alignments across the substrate-layer interface, to specific device requirements. When the strained-layer thickness exceeds the critical value, it becomes energetically favourable for the strain to relax via the introduction of dislocations at the substrate-layer interface. If the strained layer is to be used as the active region of the device then the introduction of dislocations is undesirable, since the dislocations reduce the strain in the layer (thereby altering its electronic properties) and act as carrier recombination centres, reducing carrier lifetime. Alternatively, the strained layer may be employed as a buffer layer. In this case complete relaxation of the layer

to its natural lattice constant is required. The resulting relaxed layer is then used as a substrate for the active region of the device. The purpose of the buffer layer is to provide a substrate of the desired lattice size and electronic characteristics that is integrated in a natural way with existing substrate (usually Si) technology. In this application introduction of misfit dislocations is desired in large numbers, although the presence of threading segments (segments that pass from the end of a segment of misfit dislocation at the interface up into the layer) is to be avoided, since they compromise the utility of the buffer layer as a device substrate.

Regardless of the particular application, of key importance to those in the semiconductor device industry would be the ability to predict with reliability the level of dislocation injection and consequent strain relaxation in strained-layer devices. With this in mind, specific device applications will not be considered here, except insofar as they define the geometry of strained structures. This work will then be concerned with the estimation of the extent of dislocation formation and strain relaxation in these structures. (For extensive treatises on the applications of strained-layer structures and on the effects of strain on electronic properties see [56], [8], and [39].)

Study of the process of strain relaxation may be subdivided into two parts. The first concerns the propagation and extension of dislocations already present in the layer, and the second concerns the nucleation of new dislocations. The subject matter of this thesis divides naturally into these two parts. The propagation and interaction of dislocations is considered first. A segment of misfit dislocation lying at the interface between a strained layer and its substrate terminates either where the interface meets a lateral edge of the layer, or at the boundary of the body via a so-called threading segment, which either penetrates deep into the substrate or passes upwards through the strained layer to the free surface. (In this description the simplest possible strained layer structure has been assumed, consisting of a single strained layer bounded by its substrate and a free surface. It is sometimes necessary to study strain relaxation in a buried structure, in which the strained layer is bounded on both sides by substrate material. In this case the threading segment may join segments of misfit dislocation at the two interfaces. Although differing in details, the situation is in principle very similar to the one described here.) The interfacial segment of misfit dislocation extends as the threading segment that joins it to the free surface propagates across the layer, depositing misfit dislocation in its wake. It is illustrated in Chapter 3 that the propagation and interaction of dislocations may be treated, at least approximately, using two-dimensional solutions only. A solution that proves very useful is that for the stress and energy of an array of straight misfit dislocations in a strained layer. This solution, and minor but useful modifications thereof, is therefore derived first, in Chapter

2. The layer and substrate together are modelled as a generally anisotropic half-space. Closed form solutions are obtained under the assumption that the layer and substrate have equal elastic constants. Straightforward modifications yield solutions for multiple arrays (i.e. sets of more than one array in the interface, each inclined to the others at some non-zero angles, e.g. a pair of orthogonal arrays) and for alternating arrays (i.e. arrays in which one or more component of the Burgers vector alternates in sign within the array). As a refinement, an approximate analysis is developed to assess the effect of a difference between the elastic properties of the layer and substrate. The method is based upon representing the displacement as a perturbation expansion about the solution for a homogeneous body. The correction terms, which must be added to the solutions for a homogeneous body, are obtained, accurate to first order in the difference between the elastic constants of the layer and substrate. These solutions are simply derived in Chapter 2 with their interpretation and application to the strain relaxation process being considered in Chapter 3.

In Chapter 3 the expression is presented, as derived by Freund[22], for the driving force on a dislocation tending to move its threading segment so as to extend or contract its associated misfit segment. The introduction of force concepts leads naturally to the consideration of critical values of parameters at which the forces under consideration change sign. One example of such a critical combination is the value of layer thickness, for a given lattice mismatch, at which the first dislocation propagates across an otherwise dislocation-free layer: the so-called Matthews and Blakeslee critical thickness[50]. Another approach is to deal with the total energy of a configuration containing dislocations and to define critical combinations of parameters to be those that minimize the energy. Such an approach underpins the van der Merwe definition of the critical thickness[69]. Using the solutions of Chapter 2, the Matthews and Blakeslee and van der Merwe definitions of critical thickness are compared and contrasted within a unified framework. By considering the force tending to move a single threading dislocation, the Matthews and Blakeslee view of strain relaxation is essentially *sequential*, with threading dislocations crossing the layer one at a time. In contrast, the energy-minimization approach of van der Merwe is mute concerning the particulars of dislocation introduction: it is assumed that some mechanism exists by which the lowest energy state may be attained. The two approaches lead to definitions of the critical layer thickness that may be quite different, although the differences between the numerical values obtained using these definitions are small.

Having considered the onset of the strain relaxation process in some detail, the subsequent propagation and interaction of dislocations is considered. Sequential and energy-minimization approaches to the determination of the final dislocation concentration are defined, and are

shown to lead to rather different conclusions. In general, the sequential approach is more appropriate for mechanistic treatments of strain relaxation. In particular, a phenomenon that has no obvious formulation within the energy minimization approach is the interaction of a propagating threading dislocation with a perpendicular misfit dislocation in its path. This problem has been considered by Freund[22, 23] and the approach followed here is quite similar although it differs in details. Under certain circumstances such an interaction may impede the motion of a threading dislocation to the extent of its being stopped altogether. Such events clearly have a profound effect on the strain relaxation process. The treatment described above leads to critical combinations of lattice mismatch, dislocation spacing, and layer thickness for which the motion of a threading dislocation is just halted. These critical combinations are calculated for the case of full anisotropy, using the solutions of Chapter 2.

At the end of Chapter 3 the various critical conditions derived are compared with experiment. Although agreement can be obtained in some cases, a major limitation is the lack, in the force and energy analyses presented, of any time dependent elements. Time dependence arises in two ways. Firstly, threading dislocations propagate across the layer at a finite speed, which limits the relaxation rate. Another limiting factor is the rate at which new dislocations can be nucleated. Dislocation velocity has been studied extensively and so this aspect of the time dependence is not considered here. (The reader is referred to the book by Jain[37] for a review of this work and an extensive list of references.) However, the rate at which dislocations nucleate has received much less attention, and forms the motivation for the second half of the thesis, namely Chapters 4 to 6.

Activation barriers inherent in nucleation processes are likely to be a major limiting factor in the rate at which strain relief occurs. In other words, although the energy of a given system may be lowered by the introduction of new dislocations, the mechanism by which these dislocations are formed requires that the system pass through a higher energy (or 'activated') state. The activation energy required for this process must be supplied by thermal fluctuations, and it is the rate of supply of this thermal energy that can prove to be a rate-limiting process.

Kamat and Hirth[42] have presented an Arrhenius-type expression for the steady-state loop nucleation rate, whose derivation is based on concepts developed in the study of nucleation of droplets in supersaturated vapours[18]. When a loop is nucleated in a strained layer, the elastic energy of the system initially increases with increasing loop radius, attains a maximum value, and then decreases as the loop expands further across the layer. The activation energy involved in the Arrhenius exponential in the expression of Kamat and Hirth is this maximum elastic loop energy. Clearly, a prerequisite for accurate assessment of the nucleation rate is a reliable

representation for the energy of a dislocation loop.

The elastic energy of a dislocation loop in the presence of a background stress field consists of two components: the self-energy of the loop, which is the energy of the loop in the absence of a background stress field, and the interaction energy, which is the energy change due to interaction between the background stress field and the dislocation. In Chapter 4 we present a solution that enables accurate assessment of the self-energy of a dislocation near a free surface. Specifically, we present a solution for the stresses due to an arbitrary dislocation in an isotropic, homogeneous half space. A line-integral representation for the stresses due to an arbitrary dislocation in an infinite body, due to Mura[52], has long been available. However, the proximity of a free surface is a major feature of many problems related to strained-layer structures, and the accuracy of infinite-body results in these circumstances requires investigation. In Chapter 4 we construct a line-integral representation for the stresses due to an arbitrary dislocation in an isotropic half-space. For a dislocation half-line the integral may be evaluated analytically, yielding the stresses in closed form.

In Chapter 5 dislocation nucleation is addressed directly. Section 5.2 is used to present a rigorous analysis of the self-energy of various buried loop and surface half loop configurations, using the representation for the stresses obtained in Chapter 4. The results obtained are compared with approximate expressions that are based on infinite-body results. It is found that the differences between the exact and approximate self-energies may be written as very simple correction factors. These self-energy results are used in Section 6.3 to obtain activation energies and nucleation rates for dislocation formation in a background stress field due solely to the lattice mismatch, assuming a planar substrate-layer interface . The rates obtained using the exact self-energies are found to be much larger than those obtained using previous approximations. The results predict that there should be a critical lattice mismatch at which nucleation driven solely by the mismatch-induced stress field should become significant. If the lattice mismatch is less than this critical value then nucleation can only occur at sources of high local stress, arising because of defects, impurities, and the like. The onset of dislocation nucleation driven solely by the uniform lattice mismatch (so-called homgeneous nucleation) is related to the experimentally observed change in dislocation distribution as the lattice mismatch is increased.

When the lattice mismatch is increased, it is not just the character of the dislocation distribution that is altered: the Burgers vectors of the dislocations making up the distribution are also observed to undergo a change. This change is addressed in Section 5.4. A mechanism is described by which a dislocation with Burgers vector of the type observed when the lattice mismatch is small may be converted to one with Burgers vector of the type observed when the

lattice mismatch is large. The mechanism involves a nucleation event occuring at the site of the dislocation, which changes its Burgers vector. The activation energy for this nucleation event is calculated, and is used to explain the experimental observations.

A solution with a number of applications in the study of strained semiconductor devices is provided in Chapter 6: a closed-form solution for the stresses due to a periodic array of two-dimensional mismatched inclusions in an elastic body. The solution has yet to be fully exploited, but examples are provided to illustrate its use in predicting the increased nucleation rates at interface steps and other non-planar interface features, which can arise because of substrate off-cutting or patterning. Finally, the solution is applied to the assessment of the electronic properties of strained quantum wire arrays, and used to explain the experimentally observed band gap shifts in these systems.

Closing remarks are collected in Chapter 7.

Chapter 2

Periodic Arrays of Dislocations

2.1 General Considerations

In this chapter expressions are obtained for the stresses and energy of a periodic array of straight dislocations in a generally anisotropic strained layer. With very little alteration, the solution obtained may be applied to other situations of interest: the method of solution makes consideration of multiple and alternating arrays particularly simple. The solutions are derived in this chapter without motivation. Their applications to situations of practical interest are considered in Chapter 3.

A precise description of the mathematical problem addressed is now given. An epitaxial layer is considered, whose crystal lattice fails to match that of the substrate by a mismatch strain e_{ij}^m. This is defined so that, if the layer is subjected to a strain $-e_{ij}^m$, its lattice is then brought into conformity with that of the unstressed substrate. It is supposed that the layer attempts to relax its stress by slip, giving rise to a periodic array of dislocations at the interface between the layer and substrate. The configuration is shown schematically in Fig. 2.1. The layer occupies $0 < x_2 < h$, while the underlying substrate occupies $-\infty < x_2 < 0$. The dislocations are aligned parallel to the x_3-axis and form a periodic array, with period p, along the interface. Each has Burgers vector $\mathbf{b}$; this could include a screw component (not shown explicitly in the figure).

The problem now is to estimate the stress field and energy in the composite medium, in the presence of the mismatch strain and the dislocations. It should be observed that a variety of equivalent methods is available for the treatment of dislocations in anisotropic media[63, 31, 64, 65]. However, the Fourier method used by Willis *et al.*[74] is particularly convenient for the treatment of dislocation *arrays*, and so that method is used here.

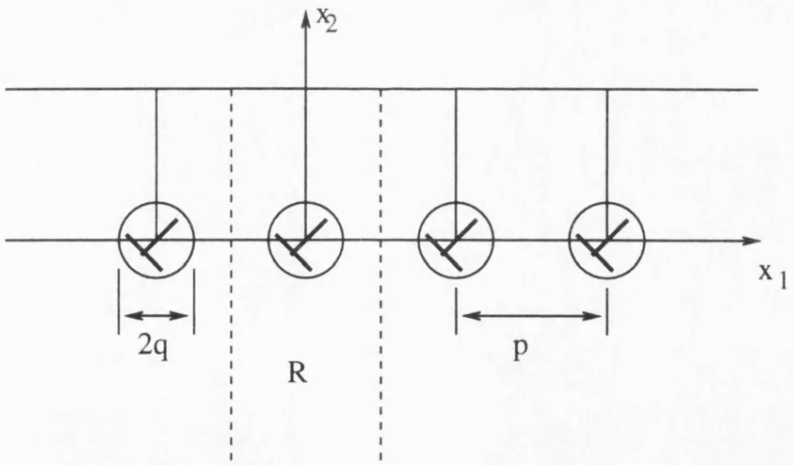

Figure 2.1: Schematic illustration of a strained-layer structure with a periodic array of dislocations lying between the layer and substrate. The surfaces across which the displacement is allowed to be discontinuous have been chosen to be perpendicular to the free surface, as illustrated. The cell R is representative of the whole structure.

Before proceeding to detailed evaluation of the stresses, we first introduce some general discussion, following Willis *et al.*, concerning evaluation of the energy. The energy of the system, which is unbounded in the x_1 and x_3 directions, is infinite, so E is defined to be the energy per unit surface area of the layer containing the dislocation array. Although the slip plane is in general inclined, it is convenient for the mathematics to take the jumps in displacement to occur across surfaces that are normal to the free surface, as shown in Fig. 2.1. Stresses are infinite at the dislocations and the usual device of introducing a core cut-off radius q will be employed: a cylindrical region, of radius q, around the dislocation core is excluded from the energy evaluation; the tractions on the surface of this cylindrical region will be taken as those necessary to maintain the usual stress field associated with the dislocation, which displays a singularity of order $1/r$ as the distance r from the dislocation tends to zero. The elastic energy in the layer, per unit length of dislocation line, is the energy in the cell denoted by R in Fig. 1, which occupies $-p/2 < x_1 < p/2$, $-\infty < x_2 < h$ and is of unit length in the x_3 direction. This energy is pE and may be written, since there is no variation of any quantity with respect to x_3,

$$pE = \frac{1}{2}\int\int_R s_{ijkl}\sigma_{ij}\sigma_{kl}dx_1dx_2, \tag{2.1}$$

where the tensor of compliances has been given as s_{ijkl} (possibly different in the layer and substrate) and σ_{ij} denotes the stress. Now, measuring displacement $\mathbf{u}$ relative to the configuration in which the layer contains no dislocations and is subjected to a strain $-e^m_{ij}$ so that it conforms

to the substrate, the stress-strain relation for the layer is

$$e_{ij} - e^m_{ij} = s_{ijkl}\sigma_{kl}. \tag{2.2}$$

The same equation can be used for the substrate if e^m_{ij} is defined to be zero there. With this convention the energy of (2.1) may be written

$$pE = \frac{1}{2}\int\int_R \sigma_{ij}(e_{ij} - e^m_{ij})dx_1 dx_2.$$

The stress field σ_{ij} is taken to be an equilibrium field, and hence divergence-free. Application of the divergence theorem to the term involving $\sigma_{ij}e_{ij}$ therefore gives

$$pE = \frac{1}{2}\int_{\partial R} \sigma_{ij}\nu_j u_i ds - \frac{1}{2}\int\int_R \sigma_{ij}e^m_{ij}dx_1 dx_2, \tag{2.3}$$

where ν is the outward normal to the domain R, the boundary of which, ∂R, is composed of the free surface, the vertical sides of the period, both sides of the surface across which $\mathbf{u}$ is allowed to jump, and the core boundary. Defining $\mathbf{b}$ to be the jump in displacement as the surface of discontinuity is crossed in the direction of (1,0,0), the periodicity of σ_{ij} and u_i, and the zero-traction condition at the free surface reduce (2.3) to

$$pE = -\frac{1}{2}\int_q^h \sigma_{i1}b_i dx_2 + \frac{1}{2}\int_{core} \sigma_{ij}\nu_j u_i ds - \frac{1}{2}\int\int_R \sigma_{ij}e^m_{ij}dx_1 dx_2. \tag{2.4}$$

If (as will be assumed) the core radius q is sufficiently small, the stress and displacement at the core boundary can be taken as those of an isolated dislocation in an infinite body. The core contribution, which we shall denote by E^c, then becomes independent of q, p, and h; this first order approximation is sufficient to ensure that the total calculated energy is independent of the orientation of the surface of discontinuity, as it must be[10]. Also, the term in (2.4) that involves e^m_{ij} may be approximated by disregarding the core entirely. Thus,

$$pE = -\frac{1}{2}\int_q^h \sigma_{i1}b_i dx_2 + E^c - \frac{1}{2}ph\bar{\sigma}_{ij}e^m_{ij}, \tag{2.5}$$

where $\bar{\sigma}_{ij}$ represents the mean stress in the layer and e^m_{ij} once more represents just the lattice mismatch between the layer and substrate.

We now proceed to evaluation of the stresses. Since the problem is linear, these may be assessed as the sum of contributions from the lattice mismatch and from the dislocations, which

will be evaluated separately.

2.2 The Stress Corresponding to the Lattice Mismatch

The stresses due to the lattice mismatch may be calculated using only simple algebraic manipulations. The lateral constraint requires that displacements can only depend on x_2; therefore, only e_{12}, e_{22}, and e_{32} can differ from zero. Also, the stress components σ_{i2} must vanish at $x_2 = h$ and are continuous across $x_2 = 0$. It is consistent to take $\sigma_{12} = \sigma_{22} = \sigma_{32} = 0$ identically. Since only e_{i2} are non-zero, (2.2) becomes

$$\sigma_{ij} = c_{ij2k}(2e_{2k} - e_{22}\delta_{2k}) - c_{ijkl}e^m_{kl}, \tag{2.6}$$

where c_{ijkl} is the tensor of stiffnesses. Setting σ_{i2} to zero we obtain

$$c_{i22k}(2e_{2k} - e_{22}\delta_{2k}) = c_{i2kl}e^m_{kl}, \tag{2.7}$$

for i =1, 2, 3. Employing the notation

$$\lambda_{ik} = c_{i22k}, \tag{2.8}$$

(2.7) may be inverted and used to write the stresses of (2.6) in terms of the mismatch as:

$$\sigma_{pq} = -[c_{pqkl} - c_{pqr2}(\lambda^{-1})_{ri}c_{i2kl})e^m_{kl}. \tag{2.9}$$

The bracketed term is the so-called planar tensor of the moduli of elasticity for the direction (0,1,0), and so (2.9) is a special case of equation (9) of the work of Orlov and Indenbom[55].

The explicit solution is particularly simple for the common case of orthorhombic symmetry with the coordinate axes taken normal to the planes of symmetry. This simplification is applicable, for example, to the study of dislocations in a cubic material, lying along a ⟨110⟩ direction in an interface with normal (100), (110), or (111). In this case all elastic constants vanish that have any index appearing an odd number of times, and (2.9) yields the following non-zero stresses:

$$\begin{aligned}
\sigma_{11} &= \frac{1}{c_{2222}}\left[\left(c^2_{1122} - c_{1111}c_{2222}\right)e^m_{11} + (c_{1122}c_{2233} - c_{1133}c_{2222})e^m_{33})\right], \\
\sigma_{13} &= -2c_{1313}e^m_{13}, \\
\sigma_{33} &= \frac{1}{c_{2222}}\left[\left(c^2_{2233} - c_{3333}c_{2222}\right)e^m_{33} + (c_{1122}c_{2233} - c_{1133}c_{2222})e^m_{11}\right].
\end{aligned} \tag{2.10}$$

2.3 The Stress Corresponding to the Dislocations

With the planes across which $\mathbf{u}$ is discontinuous chosen to be vertical, the displacement in the layer may be represented by (Willis *et al.*[74])

$$\mathbf{u} = \frac{\mathbf{b}}{2} \sum_{n=-\infty}^{\infty} \mathrm{sgn}(x_1 - np) - \frac{\mathbf{b}x_1}{p} + \frac{\mathbf{b}'x_2}{p} + \mathbf{u}^{fa}, \tag{2.11}$$

where

$$\sum_{n=-\infty}^{\infty} \mathrm{sgn}(x_1 - np) = \lim_{N\to\infty} \sum_{n=-N}^{N} \mathrm{sgn}(x_1 - np).$$

The displacement $\mathbf{u}^{fa}$ is continuous across $x_1 = np$ and periodic. The term $\mathbf{b}x_1/p$ plays the rôle of a uniform mismatch strain with components $(b_i n_j + b_j n_i)/2p$, where $\mathbf{n}$ is a unit vector parallel to $0x_1$; this is the term that tends to relax the mismatch strain e^m_{ij}. The term involving $\mathbf{b}'$ is employed to reduce the stresses σ_{i2} associated with $-\mathbf{b}x_1/p$ to zero, in the way described for e^m_{ij} in the preceding section. Following the reasoning of Section 2.2, the stresses σ_{ij} generated by the term $\mathbf{b}'x_2/p - \mathbf{b}x_1/p$ are therefore given by

$$\bar{\sigma}_{pq} = -[c_{pqkl} - c_{pqr2}(\lambda^{-1})_{ri}c_{i2kl}]\left(\frac{b_k n_l + b_l n_k}{2p}\right). \tag{2.12}$$

Again, the solution is particularly straightforward for the case of orthorhombic symmetry. Recalling that $\mathbf{n} = (1,0,0)$, it may be seen that for this symmetry (2.12) gives the following non-zero stresses:

$$\begin{aligned}
\bar{\sigma}_{11} &= \frac{1}{c_{2222}}\left(c^2_{1122} - c_{1111}c_{2222}\right)\frac{b_1}{p}, \\
\bar{\sigma}_{13} &= -c_{1313}\frac{b_3}{p}, \\
\bar{\sigma}_{33} &= \frac{1}{c_{2222}}(c_{1122}c_{2233} - c_{1133}c_{2222})\frac{b_1}{p}.
\end{aligned} \tag{2.13}$$

Reverting now to the energy, (2.5), it is convenient to represent the *total* stress σ_{ij} in the layer in the form $\sigma^{fa}_{ij} + \bar{\sigma}_{ij}$. The components of the mean stress $\bar{\sigma}_{ij}$ are the sums of the corresponding components given by (2.9) and (2.12). The stress σ^{fa}_{ij} is the fluctuating part of the array stress field (whence the 'fa' superscript), which has mean value zero in the layer and is precisely the stress associated with $\mathbf{u}^{fa}$. Equation (2.5) may thus be written

$$pE = E^{fa} - \frac{1}{2}ph\bar{\sigma}_{ij}e^r_{ij}, \tag{2.14}$$

where

$$e_{ij}^r = e_{ij}^m + \frac{b_i n_j + b_j n_i}{2p}. \tag{2.15}$$

The 'fa' superscript is used to label the energy due to the fluctuating part of the array stress field:

$$E^{fa} = -\frac{1}{2}\int_q^h \sigma_{i1}^{fa} b_i dx_2 + E^c. \tag{2.16}$$

Equation (2.14) thus represents the total energy of the system as the sum of the energy associated with a uniform relaxed mismatch strain, e_{ij}^r, and a term associated with the fluctuating strain field. Since the former strain field is constant and the latter periodic with mean value zero, these two strain fields do not interact, and hence the total energy appears as the direct sum of the energies of these components separately. Furthermore, it should be observed that the total mean stresses in the layer, $\bar{\sigma}_{ij}$, are, for *any* anisotropy, precisely the stresses obtained by making the replacement $e_{ij}^m \mapsto e_{ij}^r$ in (2.9), as is evident from inspection of (2.9), (2.12), and (2.15). It is emphasized that the term E^{fa} under no circumstances represents the energy of the dislocation array[74, 26]: part of the dislocation strain field has been absorbed into the relaxed mismatch strain.

Willis *et al.*[74] prescribed a general method by which to find the fluctuating stress and displacement fields; this involves exploiting the periodicity of these fields in x_1 by representing $\mathbf{u}^{fa}$ as a Fourier series. Its Fourier coefficients satisfy ordinary differential equations in x_2 and can be found explicitly. The Fourier coefficients of the stress components σ_{ij}^{fa} then follow, and the Fourier series for σ_{ij}^{fa}, evaluated at $x_1 = 0$, can then be integrated term by term to give a corresponding series for $E^{fa} - E^c$.

For the remainder of this chapter, summation is always indicated explicitly: the summation convention is not adopted. It is convenient to consider first a single isolated dislocation (the case of $p \to \infty$). In this case (2.11) reduces to

$$\mathbf{u}^s = \frac{1}{2}\mathbf{b}\,\mathrm{sgn}x_1 + \mathbf{u}^{fs}.$$

The 's' superscript has been introduced to emphasize that a single dislocation is being considered. The problem is solved by Fourier transforming with respect to x_1, so that

$$\tilde{\mathbf{u}}^{fs} = \int dx_1 e^{i\xi x_1} \mathbf{u}^{fs}(x_1, x_2),$$

and so on. The field $\mathbf{u}^{fs}$ obeys the equations of equilibrium. Since it is assumed that there is

no variation with respect to x_3, these transform to:

$$\sum_j \left[c_{i2j2} \frac{\partial^2 \tilde{u}_j^{fs}}{\partial x_2^2} - \mathrm{i}\xi(c_{i1j2} + c_{i2j1}) \frac{\partial \tilde{u}_j^{fs}}{\partial x_2} - \xi^2 c_{i1j1} \tilde{u}_j^{fs} \right] = 0. \tag{2.17}$$

The convention is adopted that Roman subscripts range over the values 1, 2, 3, whereas Greek subscripts range over 1, 2. Now $\mathbf{m}\mathrm{e}^{-\mathrm{i}\gamma\xi x_2}$ is a solution of (2.17) if the vector $\mathbf{m}$ satisfies

$$\mathbf{L}(1,\gamma)\mathbf{m} = 0, \tag{2.18}$$

where

$$L_{ij}(\xi_1,\xi_2) = \sum_{\alpha,\beta} c_{i\alpha j\beta}\xi_\alpha\xi_\beta. \tag{2.19}$$

Equation (2.18) only has non-trivial solutions if

$$D(1,\gamma) = 0, \tag{2.20}$$

where

$$D(\xi_1,\xi_2) = \mathrm{Det}(\mathbf{L}(\xi_1,\xi_2)). \tag{2.21}$$

In general, (2.20) is a sextic in γ whose solutions occur in complex-conjugate pairs, this last statement being a consequence of requiring that the strain-energy density be positive[17]. We define γ_r, r=1, 2, 3, to be the roots of (2.20) that have positive imaginary part, and we assume in the sequel that these are distinct. We denote by $\mathbf{m}_r$ an eigenvector of $\mathbf{L}(1,\gamma_r)$; observe that since, for each r, γ_r satisfies (2.20), a possible choice of eigenvector is

$$\mathbf{m}_r = \mathbf{N}(1,\gamma_r)\mathbf{a}_r,$$

for *any* constant vector $\mathbf{a}_r$, where

$$N_{ij}(\xi_1,\xi_2) = \{\mathrm{Adj}[\mathbf{L}(\xi_1,\xi_2)]\}_{ij}. \tag{2.22}$$

The general solution of (2.17) may thus be written:

$$\tilde{\mathbf{u}}^{fs} = \sum_r \left(\mathbf{N}_r\mathbf{a}_r^0\mathrm{e}^{-\mathrm{i}\xi\gamma_r x_2} + \overline{\mathbf{N}_r\mathbf{b}_r^0}\mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_r}x_2} \right), \tag{2.23}$$

where $\mathbf{a}_r^0$ and $\mathbf{b}_r^0$ are arbitrary constant vectors, $\mathbf{N}_r$ is used as a shorthand notation for $\mathbf{N}(1,\gamma_r)$

(this notation will be used in an analogous manner for all matrices), and a bar denotes complex conjugation. Equation (2.23) gives the solution in the layer; the requirement that the displacement be bounded as $x_2 \to -\infty$ allows the solution in the substrate to be written, for $\xi > 0$, as:

$$\tilde{\mathbf{u}}^{fs\star} = \sum_r \mathbf{N}_r \mathbf{c}_r^{0\star} \mathrm{e}^{-\mathrm{i}\xi\gamma_r x_2}. \tag{2.24}$$

The '$\star$' superscript is used to denote quantities in the substrate, and the '0' superscript emphasizes that the solution considered is for the case that the layer and substrate have identical elastic properties.

To complete the description of $\tilde{\mathbf{u}}^s$ we require the Fourier transform of $\frac{1}{2}\mathrm{sgn}x_1$, which is the generalized function $\frac{\mathrm{i}}{\xi}$ (interpreted as a Cauchy principal value).

Now, defining matrices $\mathbf{C}^{(k)}$ by

$$C_{ij}^{(k)}(\xi_1, \xi_2) = \sum_\alpha c_{ikj\alpha}\xi_\alpha, \tag{2.25}$$

and traction vectors $\mathbf{t}^{(k)}$ associated with $\mathbf{u}^{fs}$ by

$$t_i^{(k)} = \sigma_{ik}^{fs} = \sum_{j,\alpha} c_{ikj\alpha} u_{j,\alpha}^{fs}, \tag{2.26}$$

we obtain:

$$\tilde{\mathbf{t}}^{(k)} = -\mathrm{i}\xi \sum_r \left(\mathbf{C}_r^{(k)} \mathbf{N}_r \mathbf{a}_r^0 \mathrm{e}^{-\mathrm{i}\xi\gamma_r x_2} + \overline{\mathbf{C}_r^{(k)} \mathbf{N}_r \mathbf{b}_r^0} \mathrm{e}^{-\mathrm{i}\xi\overline{\gamma}_r x_2} \right) \tag{2.27}$$

in the layer, and

$$\tilde{\mathbf{t}}^{(k)\star} = -\mathrm{i}\xi \sum_r \mathbf{C}_r^{(k)} \mathbf{N}_r \mathbf{c}_r^{0\star} \mathrm{e}^{-\mathrm{i}\xi\gamma_r x_2}$$

in the substrate.

The conditions of continuity of total displacement and traction across $x_2 = 0$ require that

$$\sum_r \left[\mathbf{N}_r (\mathbf{c}_r^{0\star} - \mathbf{a}_r^0) - \overline{\mathbf{N}_r \mathbf{b}_r^0} \right] = \frac{\mathbf{i}\mathbf{b}}{\xi}, \tag{2.28}$$

and

$$\sum_r \left[\mathbf{C}_r^{(2)} \mathbf{N}_r (\mathbf{c}_r^{0\star} - \mathbf{a}_r^0) - \overline{\mathbf{C}_r^{(2)} \mathbf{N}_r \mathbf{b}_r^0} \right] = 0, \tag{2.29}$$

respectively. The zero traction condition at $x_2 = h$ requires that

$$\sum_r \left(\mathbf{C}_r^{(2)} \mathbf{N}_r \mathbf{a}_r^0 \mathrm{e}^{-\mathrm{i}\xi\gamma_r h} + \overline{\mathbf{C}_r^{(2)} \mathbf{N}_r \mathbf{b}_r^0} \mathrm{e}^{-\mathrm{i}\xi\overline{\gamma}_r h} \right) = 0. \tag{2.30}$$

With reference to the identities given in (A1) and (A2) of Appendix A it may be seen that eqns. (2.28) and (2.29) are satisfied by making the choices

$$\mathbf{c}_r^{0\star} - \mathbf{a}_r^0 = \mathbf{b}_r^0 = \frac{\mathrm{i}(\mathbf{C}_r^{(2)})^{\mathrm{T}}}{\frac{\partial D}{\partial \xi_2}(1,\gamma_r)}\frac{\mathbf{b}}{\xi}. \tag{2.31}$$

Equation (2.31) allows us enough freedom to simplify the zero traction condition by stipulating that

$$\mathbf{a}_r^0 \mathrm{e}^{-\mathrm{i}\xi\gamma_r h} = \mathbf{d}^0, \tag{2.32}$$

for all r, where $\mathbf{d}^0$ is a constant vector of our choosing. Equation (2.30) may now be inverted to yield

$$\mathbf{d}^0 = -\sum_r \mathbf{P}^{-1}\overline{\mathbf{C}_r^{(2)}\mathbf{N}_r\mathbf{b}_r^0}\mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_r}h}, \tag{2.33}$$

where

$$\mathbf{P} = \sum_r \mathbf{C}_r^{(2)}\mathbf{N}_r. \tag{2.34}$$

Equations (2.23), (2.24), (2.31), (2.32), and (2.33) completely define the Fourier transform of the displacement field in the body for $\xi > 0$. Now note that since the displacement field $\mathbf{u}^{fs}$ must be real we have

$$\overline{\tilde{\mathbf{u}}^{fs}(\xi,x_2)} = \int dx_1 \mathrm{e}^{-\mathrm{i}\xi x_1}\mathbf{u}^{fs}(x_1,x_2) = \tilde{\mathbf{u}}^{fs}(-\xi,x_2),$$

and so the Fourier transforms for $\xi < 0$ are obtained immediately.

Treatment of the periodic array of dislocations is similar to that of a single dislocation except that Fourier series are now employed. In this case we define the Fourier coefficient $\tilde{\mathbf{u}}^{fa}$ of $\mathbf{u}^{fa}$, which is the fluctuating part of the strain field due to an *array* of dislocations, to be

$$\tilde{\mathbf{u}}^{fa}(\xi,x_2) = \int_{-p/2}^{p/2} dx_1 \mathrm{e}^{\mathrm{i}\xi x_1}\mathbf{u}^{fa}(x_1,x_2),$$

where ξ takes values $2\pi k/p$, k being an integer. The displacement $\mathbf{u}^{fa}$ of (2.11) satisfies the equilibrium equations, and so its Fourier coeffiecients may be expressed in the form of (2.23) and (2.24). Since the term $-\mathbf{b}x_1/p + \mathbf{b}'x_2/p$ (see (2.11)) contributes nothing to the stresses σ_{i2}, the continuity of traction condition at $x_2 = 0$ and the zero traction condition at $x_2 = h$ are satisfied in the same way as for a single dislocation. The only new feature, in comparison with the case of the single dislocation, occurs in the satisfaction of the condition of continuity of total displacement across $x_2 = 0$, where the Fourier coefficient of $\mathbf{b}(\frac{1}{2}\mathrm{sgn}x_1 - \frac{x_1}{p})$ is needed, in place

of the Fourier transform of $\frac{1}{2}\mathbf{b}\mathrm{sgn}x_1$. The Fourier coefficient is

$$\mathbf{b}\frac{\mathrm{i}}{\xi}\left[1-\frac{2\sin(\xi p/2)}{\xi p}\right]. \tag{2.35}$$

The Fourier coefficients for the array are therefore obtained from the corresponding Fourier transforms for the single dislocation by making the replacement

$$\frac{\mathbf{b}}{\xi}\mapsto\mathbf{b}\left[\frac{1}{\xi}-\frac{2\sin(\xi p/2)}{\xi^2 p}\right].$$

Observe that the bracketed term is zero when $\xi = 0$ and $1/\xi$ when $\xi = 2\pi k/p$, k any integer different from zero. The stresses for the array are therefore obtained by summing the relevant Fourier series, omitting the term for $k = 0$; moreover, for k not equal to zero, the Fourier coefficients are given *precisely* by the corresponding Fourier transforms for the single dislocation. The stresses due to the array are therefore known through the Fourier transforms of the traction vectors $\mathbf{t}^{(k)}$. The stresses will not be dealt with directly here, as it will be found more convenient to work with them in transformed form. We now move on to the evaluation of some energy quantities.

2.4 The Energy of an Array of Dislocations

The remaining term to be evaluated in (2.14) is the energy of the fluctuating part of the strain field due to the dislocation array:

$$E^{fa}=-\frac{1}{2}\sum_i\int_q^h\sigma_{i1}^{fa}b_i dx_2+\frac{1}{2}\sum_{i,j}\int_{core}\sigma_{ij}^a\nu_j u_i^a ds. \tag{2.36}$$

We denote by E^c the second of the terms in (2.36) and concentrate initially on the evaluation of $E^{fa}-E^c$.

Recall that $\sigma_{i1}^{fa}=t_i^{(1)}$. We have, from (2.27), for the Fourier coefficient of $\mathbf{t}^{(1)}$:

$$\tilde{\mathbf{t}}^{(1)}=-\mathrm{i}\xi\sum_r\left(\mathbf{C}_r^{(1)}\mathbf{N}_r\mathbf{a}_r^0\mathrm{e}^{-\mathrm{i}\xi\gamma_r x_2}+\overline{\mathbf{C}_r^{(1)}\mathbf{N}_r\mathbf{b}_r^0}\mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_r}x_2}\right), \tag{2.37}$$

with $\mathbf{a}_r$ and $\mathbf{b}_r$ given by (2.31) and (2.32). The integral in (2.36) is evaluated at $x_1 = 0$ and so only the Fourier coefficients are of relevance. The Fourier series may be integrated term by

term and then summed, omitting the term for $k = 0$. Writing

$$-\frac{1}{2}\int_q^h \mathbf{b}^{\mathrm{T}}\mathbf{t}^{(1)}dx_2 = \frac{1}{p}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \mathbf{b}^{\mathrm{T}}\mathbf{f}^0\left(\frac{2\pi k}{p}\right) \tag{2.38}$$

we have, for $\xi > 0$, from (2.37), (2.31), and (2.32),

$$\begin{aligned}\mathbf{f}^0(\xi) &= -\frac{\mathrm{i}}{2}\left\{\sum_{r,s}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s(\mathbf{C}_s^{(1)})^{\mathrm{T}}}\left[\mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_s}h} - \mathrm{e}^{-\mathrm{i}\xi(\overline{\gamma_s}h+\gamma_r q-\gamma_r h)}\right]\right. \\ &\quad \left.-\sum_r \frac{\overline{\mathbf{C}_r^{(1)}\mathbf{N}_r\mathbf{R}_r}}{\overline{\gamma_r}}\left(\mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_r}h} - \mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_r}q}\right)\right\}\frac{\mathbf{b}}{\xi},\end{aligned} \tag{2.39}$$

where

$$\begin{aligned}\mathbf{Q}_r &= \frac{\mathbf{C}_r^{(2)}\mathbf{N}_r}{\frac{\partial D}{\partial \xi_2}(1,\gamma_r)}, \\ \mathbf{R}_r &= \frac{(\mathbf{C}_r^{(2)})^{\mathrm{T}}}{\frac{\partial D}{\partial \xi_2}(1,\gamma_r)}.\end{aligned} \tag{2.40}$$

Recalling that $\mathbf{u}(-\xi) = \overline{\mathbf{u}(\xi)}$, and therefore that $\mathbf{f}^0(-\xi) = \overline{\mathbf{f}^0(\xi)}$, we obtain:

$$E^{fa} - E^c = \frac{2}{p}\mathrm{Re}\sum_{k=1}^{\infty}\mathbf{b}^{\mathrm{T}}\mathbf{f}^0\left(\frac{2\pi k}{p}\right). \tag{2.41}$$

Each term in $\mathbf{f}^0(\xi)$, as given by (2.39), contains a factor that decays exponentially as $\xi \to \infty$, so the series of (2.41) converges. Now, for $\mathrm{Re}(\alpha) > 0$,

$$\sum_{k=1}^{\infty}\mathrm{e}^{-2\pi k\alpha/p} = \frac{\mathrm{e}^{-2\pi\alpha/p}}{1-\mathrm{e}^{-2\pi\alpha/p}} \tag{2.42}$$

gives the sum of a geometric series. Integration with respect to α yields

$$\sum_{k=1}^{\infty}\frac{p}{2\pi k}\mathrm{e}^{-2\pi k\alpha/p} = \frac{p}{2\pi}\ln\left(\frac{1}{1-\mathrm{e}^{-2\pi\alpha/p}}\right). \tag{2.43}$$

Examination of (2.39) reveals that all of the series required by (2.41) are of the form of (2.43). Completion of the algebra yields

$$E^{fa} = \mathrm{Im}\frac{\mathbf{b}^{\mathrm{T}}}{2\pi}\left\{\sum_{r,s}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s(\mathbf{C}_s^{(2)})^{\mathrm{T}}}\ln\left[\frac{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_s}-\gamma_r)h/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}\overline{\gamma_s}h/p}}\right]\right.$$

$$+\sum_r \frac{\overline{\mathbf{C}_r^{(1)}\mathbf{N}_r\mathbf{R}_r}}{\overline{\gamma_r}} \ln\left[\frac{p(1-\mathrm{e}^{-2\pi\mathrm{i}\overline{\gamma_r}h/p})}{2\pi\mathrm{i}\overline{\gamma_r}q}\right]\Bigg\}\mathbf{b} + E^c. \tag{2.44}$$

The matrices $\mathbf{N}$, $\mathbf{C}^{(k)}$, $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ are defined in (2.22), (2.25), (2.34), and (2.40). The limit of the above expression as $p \to \infty$ gives the energy of an isolated dislocation:

$$\begin{aligned} E^s &= \mathrm{Im}\frac{\mathbf{b}^{\mathrm{T}}}{2\pi}\Bigg\{\sum_{r,s}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s(\mathbf{C}_s^{(2)})^{\mathrm{T}}}\ln\left[\frac{\overline{\gamma_s}-\gamma_r}{\overline{\gamma_s}}\right] \\ &\quad +\sum_r \frac{\overline{\mathbf{C}_r^{(1)}\mathbf{N}_r\mathbf{R}_r}}{\overline{\gamma_r}}\ln\left(\frac{h}{q}\right)\Bigg\}\mathbf{b} + E^c. \end{aligned} \tag{2.45}$$

Finally, we discuss evaluation of the term E^c, which is given by:

$$E^c = \frac{1}{2}\int_{core}\sigma_{ij}^a\nu_j u_i^a ds.$$

This is simplified by approximating the field close to a dislocation in the array by the field of an isolated dislocation in an infinite body. The error incurred in this approximation is of the order of q/h, q/p, which must be insignificant in any case if linear elasticity theory is to give an accurate estimate of the energy. In an infinite body the zero traction condition at $x_2 = h$ is replaced by the requirement that the displacement be bounded as $x_2 \to \infty$, which is fulfilled by choosing $\mathbf{a}_r^0 = 0$. The constant vectors $\mathbf{b}_r^0$ and $\mathbf{c}_r^{0\star}$ are still defined by (2.31), and the Fourier coefficients for the displacement and stresses are defined by (2.23) and (2.27), together with the corresponding equations for the substrate. The stress and displacement fields for an array may be found by summing the relevant Fourier series, and the fields for an isolated dislocation recovered by letting $p \to \infty$. Even with this approximation, E^c must be evaluated numerically. For a detailed acount of this procedure the reader is referred to the work of Bacon, Bullough, and Willis[3].

2.5 Multiple Arrays

The given decomposition of the total energy, (2.14), into the sum of the energy due to a relaxed mismatch strain and that due to a fluctuating strain field that has mean value zero is particularly convenient when considering more than one array. Suppose, for example, that two dislocation arrays A and $\hat{A}$, of periods p and $\hat{p}$ respectively, were present in the interface and inclined to one another at some non-zero angle. The stress σ_{ij}^{fa} due to one would vary periodically along the length of a dislocation in the other array, and so there would be no interaction of the energy terms

E^{fa} and $\hat{E}^{fa}$. Only the mean stresses would contribute to the interaction between the energies of the two arrays. But this interaction is accounted for precisely in the energy contribution from the relaxed mismatch strain, which is now relaxed by contributions from each array. The total energy per unit area in this case can be written

$$E = \frac{1}{p}E^{fa} + \frac{1}{\hat{p}}\hat{E}^{fa} - \frac{1}{2}h\bar{\sigma}_{ij}e^r_{ij}, \tag{2.46}$$

where now

$$e^r_{ij} = e^m_{ij} + \frac{b_i n_j + b_j n_i}{2p} + \frac{\hat{b}_i \hat{n}_j + \hat{b}_j \hat{n}_i}{2\hat{p}}. \tag{2.47}$$

Here $\hat{\mathbf{b}}$ and $\hat{\mathbf{n}}$ represent the Burgers vector and normal for the array $\hat{A}$ (see discussion after (2.11)). In an anisotropic medium E^{fa} will in general depend upon the orientation of the dislocation line. The term $\hat{E}^{fa}$ must therefore be calculated using elastic constants referred to a coordinate system in which the x_3 axis is aligned in the direction of a dislocation in the array $\hat{A}$. As before, the mean stress $\bar{\sigma}_{ij}$ appearing in (2.46) is obtained by making the replacement $e^m_{ij} \mapsto e^r_{ij}$ in (2.9). The same procedure may be extended to consider any number of inclined arrays.

2.6 Alternating Arrays

It might be supposed that under certain circumstances the energy of an array would be lowered by alternating within the array the sign of one or more of the components of the Burgers vector, for example the screw component. The energy of such an alternating array is now calculated.

Consider an array, with inter-dislocation spacing p, in which dislocations at positions $x_1 = 2np$ have Burgers vector $\mathbf{b}$ and those in positions $x_1 = (2n+1)p$ have Burgers vector $\tilde{\mathbf{b}}$. The array may be considered to be the superposition of two arrays of spacing $2p$, with dislocations in one array having Burgers vector $\mathbf{b}$ and those in the other array having Burgers vector $\tilde{\mathbf{b}}$. The energy may then be written as the sum of the energies of the two arrays separately, which are known from Section 2.4, together with an interaction term that remains to be evaluated.

Consider, then, an alternating array, so that

$$\tilde{\mathbf{b}} = (\tilde{b}_1, \tilde{b}_2, \tilde{b}_3) = (\pm b_1, \pm b_2, \pm b_3),$$

where signs are *not* ordered. As in the case of the non-alternating array the strain field may be split into a uniform part and a fluctuating part that has mean value zero. The energy

contribution from the uniform part of the strain field may be obtained from the corresponding energy contribution in the non-alternating case by setting to zero all components of the Burgers vector that alternate in sign within the array. This prescription accounts precisely for the interaction between the uniform parts of the strain fields of the arrays with Burgers vectors $\mathbf{b}$ and $\tilde{\mathbf{b}}$.

Evaluation of the energy due to the fluctuating part of the strain field is more complicated. We chose to evaluate first the energy associated with the fluctuating part of the strain field due to an alternating array with inter-dislocation spacing $p/2$; the required result for an array with inter-dislocation spacing p is then trivially obtained. The dislocations with Burgers vector $\mathbf{b}$ now occupy positions $x_1 = np$, and those with Burgers vector $\tilde{\mathbf{b}}$ are at $x_1 = (n + \frac{1}{2})p$. Appropriate use of the divergence theorem, in a manner similar to that of Section 2.1 (see, for example, Appendices A and B of [25]), allows the energy per unit length of dislocation due to the fluctuating part of the strain field in an alternating array with inter-dislocation spacing $p/2$ to be written

$$\frac{1}{2}\left[E^{fa}(\mathbf{b}) + E^{fa}(\tilde{\mathbf{b}})\right] + E^{i},$$

where the interaction energy E^i is given by

$$E^i = -\frac{1}{2}\int_0^h \sigma_{i1}^{fa}(\mathbf{b})\tilde{b}_i dx_2. \tag{2.48}$$

Here $\sigma_{i1}^{fa}(\mathbf{b})$ is the fluctuating part of the stress field of an array of dislocations of Burgers vector $\mathbf{b}$ and period p. The integral in (2.48) is evaluated at $x_1 = p/2$, half-way between two dislocations with Burgers vector $\mathbf{b}$, i.e. on the surface of discontinuity of a dislocation with Burgers vector $\tilde{\mathbf{b}}$. The interaction energy may be written

$$E^i = -\frac{1}{2}\int_0^h \tilde{\mathbf{b}}^{\mathrm{T}}\mathbf{t}^{(1)}(\mathbf{b})dx_2,$$

where the integral is evaluated at $x_1 = p/2$. The Fourier coefficients of $\mathbf{t}^{(1)}$ are known through (2.27). Using the Fourier series representation of $\mathbf{t}^{(1)}$ we may write

$$-\frac{1}{2}\int_0^h \tilde{\mathbf{b}}^{\mathrm{T}}\mathbf{t}^{(1)}(\mathbf{b})dx_2 = \frac{1}{p}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \tilde{\mathbf{b}}^{\mathrm{T}}\left[-\frac{1}{2}\int_0^h \tilde{\mathbf{t}}^{(1)}\left(\frac{2\pi k}{p}\right)dx_2\right] \mathrm{e}^{-2\pi \mathrm{i} k x_1/p}.$$

The term in square brackets is known from (2.38) and (2.39) by taking the case $q = 0$. Evaluating

at $x_1 = p/2$ and using the fact that $\hat{\mathbf{u}}(-\xi) = \overline{\hat{\mathbf{u}}(\xi)}$ we obtain

$$E^i = \frac{2}{p}\mathrm{Re}\sum_{k=1}^{\infty}\tilde{\mathbf{b}}^{\mathrm{T}}(-1)^k\mathbf{g}^0\left(\frac{2\pi k}{p}\right), \tag{2.49}$$

where $\mathbf{g}^0(\xi)$ is obtained from $\mathbf{f}^0(\xi)$ (see (2.39)) by putting $q = 0$ there. The series required by (2.49) is thus

$$\sum_{k=1}^{\infty}(-1)^k\frac{p}{2\pi k}\mathrm{e}^{-2\pi k\alpha/p} = \frac{p}{2\pi}\ln\left(\frac{1}{1+\mathrm{e}^{-2\pi\alpha/p}}\right). \tag{2.50}$$

Completion of some elementary algebra, followed by the replacement $p/2 \mapsto p$, yields the energy E^{fa}_{alt} due to the fluctuating part of the strain field of an alternating array with spacing p:

$$E^{fa}_{alt} = \frac{1}{2}\left[E^{fa}(\mathbf{b}, 2p) + E^{fa}(\tilde{\mathbf{b}}, 2p)\right] + E^i, \tag{2.51}$$

where

$$\begin{aligned} E^i \;=\; & \mathrm{Im}\frac{\tilde{\mathbf{b}}^{\mathrm{T}}}{2\pi}\left\{\sum_{r,s}\frac{\mathbf{C}^{(1)}_r\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s(\mathbf{C}^{(2)}_s)^{\mathrm{T}}}\ln\left[\frac{1+\mathrm{e}^{-\pi\mathrm{i}(\overline{\gamma_s}-\gamma_r)h/p}}{1+\mathrm{e}^{-\pi\mathrm{i}\overline{\gamma_s}h/p}}\right]\right. \\ & \left.+\sum_r\frac{\overline{\mathbf{C}^{(1)}_r\mathbf{N}_r\mathbf{R}_r}}{\overline{\gamma_r}}\ln\left(\frac{1+\mathrm{e}^{-\pi\mathrm{i}\overline{\gamma_r}h/p}}{2}\right)\right\}\mathbf{b}. \end{aligned} \tag{2.52}$$

It is interesting to observe that once the energy of a non-alternating array has been determined, the energy of an alternating array follows with very little further calculation. In particular, all matrices and constants appearing in (2.52) have already appeared in (2.44).

2.7 Simplifications

Certain systems of interest have symmetry properties that simplify evaluation of the energy expressions given above. In particular, orthorhombic symmetry, with each coordinate plane a plane of symmetry, results in significant simplifications. This symmetry is possessed by a large number of systems of interest; for example, a cubic material, with interface oriented in the (100), (110), or (111) direction with dislocation lines running in a ⟨110⟩ direction has the appropriate symmetry. In these cases, all elastic constants vanish that have any index appearing an odd number of times. This result was used in Sections 2.2 and 2.3 to simplify the expressions for the mean stresses in the layer. Also simplified are the matrices appearing in the energy solution, and calculation of the γ_r, which are roots of (2.20). In the case of orthorhombic symmetry,

(2.20) degenerates to

$$(c_{2323}\gamma^2 + c_{1313})\left[c_{1212}c_{2222}\gamma^4 + (c_{1111}c_{2222} - c_{1122}^2 - 2c_{1122}c_{1212})\gamma^2 + c_{1111}c_{1212}\right] = 0. \quad (2.53)$$

The problem of determining the γ_r is thus reduced to that of solving a quadratic equation. It should be stressed that the elastic constants appearing throughout this chapter are referred to the coordinate system illustrated in Fig. 1, which is defined by the dislocation line direction, and which is, in general, different from the crystallographic coordinate system.

2.8 Layer and Substrate with Different Elastic Properties

We now consider the stress field and energy of an array of dislocations at the interface between a layer and substrate that have different elastic constants. Using a '$\star$' superscript to denote quantities in the substrate we suppose that

$$c^\star_{ijkl} = c_{ijkl} + \delta_{ijkl}$$

for some tensor δ_{ijkl}. The case $\delta_{ijkl} = 0$ was solved in Sections 2.3 and 2.4. We define quantities $D^\star$, $\mathbf{N}^\star$, and $\mathbf{C}^{\star(k)}$ by replacing c_{ijkl} by $c^\star_{ijkl}$ in (2.21), (2.22), and (2.25). Moreover, the roots of

$$D^\star(1,\gamma) = 0$$

that have positive imaginary part are denoted $\gamma^\star_r$, $r = 1,2,3$. As in previous sections, the shorthand notation $\mathbf{N}^\star_r$ is used for $\mathbf{N}^\star(1,\gamma^\star_r)$, and so on.

We follow the procedure of Section 2.3 by considering first an isolated dislocation. The continuous part, $\mathbf{u}^{fs}$, of the displacement field of a single dislocation obeys the equations of equilibrium, and so its transform obeys (2.17). However, equations (2.17) are no longer the same in the layer and substrate: in the latter case the constants $c^\star_{ijkl}$ replace the constants c_{ijkl}. The solution for $\xi > 0$ is now written

$$\tilde{\mathbf{u}}^{fs} = \sum_r \left(\mathbf{N}_r\mathbf{a}_r e^{-i\xi\gamma_r x_2} + \overline{\mathbf{N}_r\mathbf{b}_r} e^{-i\xi\overline{\gamma_r} x_2}\right) \quad (2.54)$$

in the layer and

$$\tilde{\mathbf{u}}^{fs\star} = \sum_r \mathbf{N}^\star_r\mathbf{c}^\star_r e^{-i\xi\gamma^\star_r x_2} \quad (2.55)$$

in the substrate. The constant vectors $\mathbf{a}_r$, $\mathbf{b}_r$, and $\mathbf{c}^\star_r$ are determined by the conditions of

continuity of total displacement and traction across $x_2 = 0$ and the zero traction condition at $x_2 = h$.

In general, for $\delta_{ijkl} \neq 0$, the constant vectors appearing in (2.54) and (2.55) cannot be determined explicitly. We therefore assume that we may express them as perturbation expansions about the solution for $\delta_{ijkl} = 0$:

$$\begin{aligned} \mathbf{a}_r &= \mathbf{a}_r^0 + \mathbf{a}_r^1 + \mathbf{a}_r^2 + ..., \\ \mathbf{b}_r &= \mathbf{b}_r^0 + \mathbf{b}_r^1 + \mathbf{b}_r^2 + ..., \\ \mathbf{c}_r^\star &= \mathbf{c}_r^{0\star} + \mathbf{c}_r^{1\star} + \mathbf{c}_r^{2\star} + ..., \end{aligned}$$

where $\mathbf{a}_r^0$, $\mathbf{b}_r^0$, and $\mathbf{c}_r^{0\star}$ are defined through (2.31) and (2.32), and where $\mathbf{a}_r^k$ is of k^{th} order in the modulus of the perturbation, δ_{ijkl}, of the elastic constants, and so on. Further, we define matrices

$$\begin{aligned} \mathbf{K}_r &= \mathbf{N}_r^\star - \mathbf{N}_r, \\ \mathbf{J}_r &= \mathbf{C}_r^{\star(2)} - \mathbf{C}_r^{(2)}, \end{aligned} \tag{2.56}$$

which are of first order in the perturbation δ_{ijkl}. Writing the conditions of continuity of total displacement and of traction across $x_2 = 0$ and neglecting terms of second order in the perturbation, we obtain

$$\sum_r \left[\mathbf{N}_r(\mathbf{c}_r^{1\star} - \mathbf{a}_r^1) - \overline{\mathbf{N}_r \mathbf{b}_r^1}\right] = \boldsymbol{\phi}, \tag{2.57}$$

and

$$\sum_r \left[\mathbf{C}_r^{(2)}\mathbf{N}_r(\mathbf{c}_r^{1\star} - \mathbf{a}_r^1) - \overline{\mathbf{C}_r^{(2)}\mathbf{N}_r \mathbf{b}_r^1}\right] = \boldsymbol{\psi}, \tag{2.58}$$

where

$$\boldsymbol{\phi} = -\sum_r \mathbf{K}_r(\mathbf{a}_r^0 + \mathbf{b}_r^0),$$

and

$$\boldsymbol{\psi} = -\sum_r (\mathbf{C}_r^{(2)}\mathbf{K}_r + \mathbf{J}_r\mathbf{N}_r)(\mathbf{a}_r^0 + \mathbf{b}_r^0).$$

The zero traction condition at $x_2 = h$ yields

$$\sum_r \left(\mathbf{C}_r^{(2)}\mathbf{N}_r\mathbf{a}_r^1 e^{-i\xi\gamma_r h} + \overline{\mathbf{C}_r^{(2)}\mathbf{N}_r\mathbf{b}_r^1} e^{-i\xi\overline{\gamma_r} h}\right) = 0. \tag{2.59}$$

In writing (2.57), (2.58), and (2.59), use has been made of the fact that the vectors $\mathbf{a}_r^0$, $\mathbf{b}_r^0$, and $\mathbf{c}_r^{0\star}$ satisfy (2.28), (2.29), and (2.30). With reference to the identities given in (A1) through (A4)

of Appendix A it may be seen that (2.57) and (2.58) are satisfied by making the choices

$$\mathbf{c}_r^{1\star} - \mathbf{a}_r^1 = \frac{(\mathbf{C}_r^{(2)})^{\mathrm{T}}\boldsymbol{\phi} + \boldsymbol{\psi}}{\frac{\partial D}{\partial \xi_2} D(1, \gamma_r)}, \tag{2.60}$$

$$\mathbf{b}_r^1 = -\frac{(\mathbf{C}_r^{(2)})^{\mathrm{T}}\overline{\boldsymbol{\phi}} + \overline{\boldsymbol{\psi}}}{\frac{\partial D}{\partial \xi_2}(1, \gamma_r)}. \tag{2.61}$$

Following the procedure of Section 2.3, the zero traction condition is simplified by stipulating that

$$\mathbf{a}_r^1 \mathrm{e}^{-\mathrm{i}\xi\gamma_r h} = \mathbf{d}^1$$

for all r, where $\mathbf{d}^1$ is a constant vector of our choosing. Equation (2.59) may now be inverted to yield

$$\mathbf{d}^1 = -\sum_r \mathbf{P}^{-1}\overline{\mathbf{C}_r^{(2)}\mathbf{N}_r\mathbf{b}_r^1}\mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_r}h},$$

where the matrix $\mathbf{P}$ is defined in (2.34), and $\mathbf{b}_r^1$ is given in (2.61). The constant vectors $\mathbf{a}_r^1$, $\mathbf{b}_r^1$, and $\mathbf{c}_r^{1\star}$, which define the first order correction term in the perturbation expansion, have now been determined. Calculation of the Fourier coefficients for the periodic array now follows the procedure of Section 2.3. In summing the appropriate series it should be noted that the vectors $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ depend upon ξ, the variable over which summation is performed. After some elementary algebra it is found that the following first order correction term must be added to the function $\mathbf{f}^0(\xi)$ of (2.39):

$$\begin{aligned}
\mathbf{f}^1(\xi) \;=\; & -\frac{\mathrm{i}}{2}\Bigg[\sum_{r,s}\frac{\overline{\mathbf{C}_r^{(1)}\mathbf{S}_r}}{\overline{\gamma_r}}\overline{((\mathbf{C}_r^{(2)})^{\mathrm{T}}\mathbf{K}_s + \mathbf{T}_s)\mathbf{R}_s}\frac{\mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_r}x_2}}{\xi} \\
& + \sum_{r,s,t}\frac{\overline{\mathbf{C}_r^{(1)}\mathbf{S}_r}}{\overline{\gamma_r}}\overline{((\mathbf{C}_r^{(2)})^{\mathrm{T}}\mathbf{K}_s + \mathbf{T}_s)\mathbf{P}^{-1}\mathbf{Q}_t(\mathbf{C}_t^{(2)})^{\mathrm{T}}}\frac{\mathrm{e}^{-\mathrm{i}\xi(\overline{\gamma_r}x_2 + \overline{\gamma_t}h - \gamma_s h)}}{\xi} \\
& - \sum_{r,s,t}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s}\overline{((\mathbf{C}_s^{(2)})^{\mathrm{T}}\mathbf{K}_t + \mathbf{T}_t)\mathbf{R}_t}\frac{\mathrm{e}^{-\mathrm{i}\xi(\overline{\gamma_s}h + \gamma_r x_2 - \gamma_r h)}}{\xi} \\
& - \sum_{q,r,s,t}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s}\overline{((\mathbf{C}_s^{(2)})^{\mathrm{T}}\mathbf{K}_t + \mathbf{T}_t)\mathbf{P}^{-1}\mathbf{Q}_q(\mathbf{C}_q^{(2)})^{\mathrm{T}}}\frac{\mathrm{e}^{-\mathrm{i}\xi(\overline{\gamma_s}h + \gamma_r x_2 - \gamma_r h + \overline{\gamma_q}h - \gamma_t h)}}{\xi}\Bigg]\mathbf{b}\Bigg|_{x_2=q}^{h}
\end{aligned}$$

Here

$$\mathbf{S}_r = \frac{\mathbf{N}_r}{\frac{\partial D}{\partial \xi_2}(1, \gamma_r)},$$

$$\mathbf{T}_r = \mathbf{C}_r^{(2)}\mathbf{K}_r + \mathbf{J}_r\mathbf{N}_r. \tag{2.62}$$

Summation of the appropriate series, as in Section 2.4, gives the first order correction $E^{fa1} - E^{c1}$ to the energy of the fluctuating part of the strain field. Completion of the algebra yields the following result for the first order correction term that must be added to E^{fa} to account for the difference between the elastic properties of the layer and substrate:

$$\begin{aligned} E^{fa1} \;=\; & \mathrm{Im}\frac{\mathbf{b}^{\mathrm{T}}}{2\pi}\left\{\sum_{r,s}\frac{\overline{\mathbf{C}_r^{(1)}\mathbf{S}_r}}{\overline{\gamma_r}}(\overline{(\mathbf{C}_r^{(2)})^{\mathrm{T}}}\mathbf{K}_s+\mathbf{T}_s)\mathbf{R}_s\ln\left[\frac{2\pi\mathrm{i}\overline{\gamma_r}q}{p(1-\mathrm{e}^{-2\pi\mathrm{i}\overline{\gamma_r}h/p})}\right]\right. \\ & +\sum_{r,s,t}\frac{\overline{\mathbf{C}_r^{(1)}\mathbf{S}_r}}{\overline{\gamma_r}}(\overline{(\mathbf{C}_r^{(2)})^{\mathrm{T}}}\mathbf{K}_s+\mathbf{T}_s)\mathbf{P}^{-1}\overline{\mathbf{Q}_t(\mathbf{C}_t^{(2)})^{\mathrm{T}}}\ln\left[\frac{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_t}-\gamma_s)h/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_r}+\overline{\gamma_t}-\gamma_s)h/p}}\right] \\ & +\sum_{r,s,t}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s}(\overline{(\mathbf{C}_s^{(2)})^{\mathrm{T}}}\mathbf{K}_t+\mathbf{T}_t)\mathbf{R}_t\ln\left[\frac{1-\mathrm{e}^{-2\pi\mathrm{i}\overline{\gamma_s}h/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_s}-\gamma_r)h/p}}\right] \\ & \left.+\sum_{q,r,s,t}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s}(\overline{(\mathbf{C}_s^{(2)})^{\mathrm{T}}}\mathbf{K}_t+\mathbf{T}_t)\mathbf{P}^{-1}\overline{\mathbf{Q}_q(\mathbf{C}_q^{(2)})^{\mathrm{T}}}\ln\left[\frac{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_s}+\overline{\gamma_q}-\gamma_t)h/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_s}-\gamma_r+\overline{\gamma_q}-\gamma_t)h/p}}\right]\right\}\mathbf{b} \\ & +E^{c1}. \end{aligned} \tag{2.63}$$

The matrices $\mathbf{C}^{(k)}$, $\mathbf{N}$, $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{K}$, $\mathbf{S}$, and $\mathbf{T}$ are defined in (2.22), (2.25), (2.34), (2.40), (2.56), and (2.62). The uniform part of the strain field is unaffected by changing the elastic constants of the substrate.

We now require the first order correction to the core contribution, given by

$$E^{c1} = \frac{1}{2}\int_{core}(\sigma_{ij}^{a0}u_i^{a1}+\sigma_{ij}^{a1}u_i^{a0})\nu_j ds. \tag{2.64}$$

Here the '0' superscript refers to the solution for $\delta_{ijkl} = 0$, and the '1' superscript refers to the correction terms that are of first order in the perturbation. As before, the problem may be simplified by approximating the required fields by those of an isolated dislocation in an infinite body. The requirement that the displacement remain bounded as $x_2 \to \infty$ is fulfilled by choosing $\mathbf{a}_r^0 = \mathbf{a}_r^1 = 0$. The constant vectors $\mathbf{b}_r^0$, $\mathbf{c}_r^{0\star}$, $\mathbf{b}_r^1$, and $\mathbf{c}_r^{1\star}$ are still defined by (2.31), (2.60), and (2.61). The fields for the array may be found by summing the relevant Fourier series, as before, and the fields for an isolated dislocation recovered by letting $p \to \infty$. Numerical integration as required by (2.64) yields E^{c1}, thereby completing the solution.

The first order correction to the energy for alternating arrays, E^{i1}, is obtained in a similar manner to E^i (see Section 2.6) except that the first order corrections to the traction are used; in (2.49) the function $\mathbf{g}^0(\xi)$ must be replaced by a function $\mathbf{g}^1(\xi)$. This latter function is obtained by putting $q = 0$ in $\mathbf{f}^1(\xi)$ above. Completion of the algebra and the replacement $p/2 \mapsto p$ yields

the result:

$$
\begin{aligned}
E^{i1} \;=\; & \operatorname{Im}\frac{\mathbf{b}^{\mathrm{T}}}{2\pi}\Bigg\{\sum_{r,s}\frac{\overline{\mathbf{C}_r^{(1)}}\overline{\mathbf{S}_r}}{\overline{\gamma_r}}(\overline{(\mathbf{C}_r^{(2)})^{\mathrm{T}}}\mathbf{K}_s+\mathbf{T}_s)\mathbf{R}_s\ln\left[\frac{2}{1+\mathrm{e}^{-\pi\mathrm{i}\overline{\gamma_r}h/p}}\right] \\
& +\sum_{r,s,t}\frac{\overline{\mathbf{C}_r^{(1)}}\overline{\mathbf{S}_r}}{\overline{\gamma_r}}(\overline{(\mathbf{C}_r^{(2)})^{\mathrm{T}}}\mathbf{K}_s+\mathbf{T}_s)\mathbf{P}^{-1}\overline{\mathbf{Q}_t(\mathbf{C}_t^{(2)})^{\mathrm{T}}}\ln\left[\frac{1+\mathrm{e}^{-\pi\mathrm{i}(\overline{\gamma_t}-\gamma_s)h/p}}{1+\mathrm{e}^{-\pi\mathrm{i}(\overline{\gamma_r}+\overline{\gamma_t}-\gamma_s)h/p}}\right] \\
& +\sum_{r,s,t}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s}(\overline{(\mathbf{C}_s^{(2)})^{\mathrm{T}}}\mathbf{K}_t+\mathbf{T}_t)\mathbf{R}_t\ln\left[\frac{1+\mathrm{e}^{-\pi\mathrm{i}\overline{\gamma_s}h/p}}{1+\mathrm{e}^{-\pi\mathrm{i}(\overline{\gamma_s}-\gamma_r)h/p}}\right] \\
& +\sum_{q,r,s,t}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s}(\overline{(\mathbf{C}_s^{(2)})^{\mathrm{T}}}\mathbf{K}_t+\mathbf{T}_t)\mathbf{P}^{-1}\overline{\mathbf{Q}_q(\mathbf{C}_q^{(2)})^{\mathrm{T}}}\ln\left[\frac{1+\mathrm{e}^{-\pi\mathrm{i}(\overline{\gamma_s}+\overline{\gamma_q}-\gamma_t)h/p}}{1+\mathrm{e}^{-\pi\mathrm{i}(\overline{\gamma_s}-\gamma_r+\overline{\gamma_q}-\gamma_t)h/p}}\right]\Bigg\}\mathbf{b}.
\end{aligned}
$$

Chapter 3

Force and Energy Considerations

3.1 The Generalized Driving Force on A Threading Dislocation

Consider a dislocation whose threading segment is propagating across a strained layer, depositing a segment of misfit dislocation in its wake, as illustrated in Fig. 3.1. Coordinates are chosen with x_3 parallel to the misfit segment of the propagating dislocation and x_2 normal to the free surface. For later convenience an additional coordinate, η, is defined that is perpendicular to $0x_3$ and that lies in the glide plane of the propagating dislocation as illustrated. The free surface of the layer is at $x_2 = h$, and the substrate-layer interface is at $x_2 = 0$, as before. The glide plane of the dislocation is inclined at an angle α to the x_2-axis. The threading segment of the propagating dislocation depicted in the figure is depositing misfit dislocation at a height y above the interface, and its position is described by the position $(x_1, x_2, x_3) = (x, y, z) = \mathbf{x}$ of the bottom of the threading arm. We now present an expression, as derived by Freund[22], for the driving force on the threading arm of this dislocation, tending to move it so as to extend its misfit segment. This expression will then be used to consider various characteristics of the strain relaxation process.

Let the background stress field in which the dislocation moves be σ^B. This may include contributions from an applied mismatch stress, and from other dislocations already present in the layer. A prerequisite for the representation of the force is an expression for the energy change on introducing the dislocation under consideration into the stress field σ^B. We now calculate this energy change. Let S be the surface across which the displacement $\mathbf{u}$ due to the dislocation is allowed to be discontinuous. This may be chosen to be any surface bounded by the dislocation,

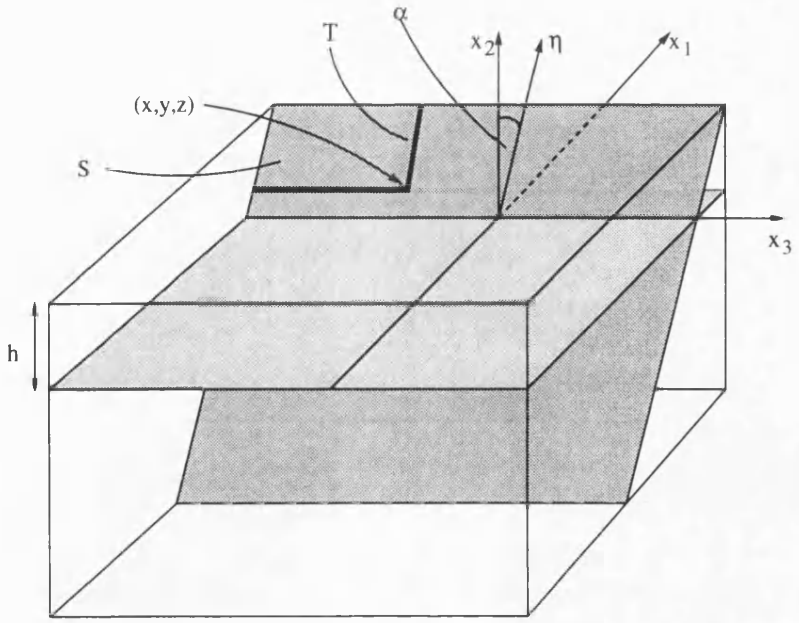

Figure 3.1: Schematic illustration of a propagating dislocation whose threading segment is depositing misfit dislocation as it moves.

but for convenience, we take it to be the section of the glide plane of the dislocation bounded by the dislocation and the free surface, as illustrated in Fig. 3.1. Define the normal $\mathbf{m}$ to the surface S and the Burgers vector $\mathbf{b}$ of the dislocation so that $\mathbf{b}$ is the jump in displacement as S is crossed in the direction of $\mathbf{m}$. We shall always choose coordinates so that the propagating dislocation moves parallel to the x_3-axis. Hence,

$$\mathbf{m} = (m_1, m_2, 0).$$

Equivalently, the Burgers vector may be defined by assigning a positive direction to the dislocation line. The direction of the normal $\mathbf{m}$ is then defined by the line direction via the usual righthand screw rule. The Burgers vector is then related to the normal as just described. (This corresponds to the F/S righthand screw convention[31].) Let the stress field of the dislocation, in the absence of a background stress field, be σ_{ij}^D. Now consider the total energy of the system, which occupies a region V, say:

$$\begin{aligned} E &= \frac{1}{2}\int_V dV(\sigma_{ij}^B + \sigma_{ij}^D)(e_{ij}^B + e_{ij}^D) \\ &= \frac{1}{2}\int_V dV \sigma_{ij}^B e_{ij}^B + \frac{1}{2}\int_V dV(\sigma_{ij}^D + 2\sigma_{ij}^B)e_{ij}^D, \end{aligned}$$

where the second equality follows from reciprocity: $\sigma^D_{ij}e^B_{ij} = \sigma^B_{ij}e^D_{ij}$. Throughout this chapter, the summation convention is adopted unless otherwise indicated. The first integral on the righthand side of the second of the above equalities is the energy of the system prior to the introduction of the dislocation under consideration. The energy change on introducing the dislocation is therefore given by

$$W = \frac{1}{2}\int_V dV(\sigma^D_{ij} + 2\sigma^B_{ij})u^D_{i,j}. \tag{3.1}$$

The fields σ^D and σ^B are assumed to be equilibrium fields, and hence divergence free. Therefore (3.1) becomes

$$\begin{aligned} W &= \frac{1}{2}\int_V dV[(\sigma^D_{ij} + 2\sigma^B_{ij})u^D_i]_{,j} \\ &= \frac{1}{2}\int_{\partial V} dS(\sigma^D_{ij} + 2\sigma^D_{ij})\nu_j u_i, \end{aligned} \tag{3.2}$$

where the second equality is obtained via the divergence theorem, noting that ν is the *outward* normal to the domain V. Now write $S^{\pm}$ for the two sides of the surface of discontinuity, S (defined so that one goes from S^- to S^+ in crossing the surface S in the direction of $\mathbf{m}$). The boundary of V, ∂V, includes S^+ and S^-, in addition to the free surface boundary of the body. We must also include a cylindrical surface, S_c of radius q, say, that surrounds the dislocation core, thus circumventing the problem of the stresses becoming singular at the core of an ideal elastic dislocation. The core radius q must be large enough so that Hooke's law is nowhere grossly violated outside S_c, and hence that linear elasticity theory, leading to (3.2), may be applied. If this surface S_c is incorporated into ∂V then (3.2) gives precisely the energy in the body, excluding the region within S_c, to account for which, some non-linear core energy may be added. Because, by definition, $\sigma_{ij}\nu_j$ is zero at a free surface, (3.2) becomes

$$W = \frac{1}{2}\int_{S^+}(\sigma^D_{ij} + 2\sigma^B_{ij})u^D_i\nu_j dS + \frac{1}{2}\int_{S^-}(\sigma^D_{ij} + 2\sigma^B_{ij})u^D_i\nu_j dS + E^c, \tag{3.3}$$

where

$$E^c = \frac{1}{2}\int_{S_c}(\sigma^D_{ij} + 2\sigma^B_{ij})u^D_i\nu_j dS.$$

If the region V is modelled as being of infinite extent in some direction, then decay of the displacement $\mathbf{u}^D$ at infinity ensures that the boundary terms vanish. Recalling that the normal ν is *outward* to the domain V, it is clear that $\nu_j = m_j$ on S^- and $\nu_j = -m_j$ on S^+. Moreover, if $u^{\pm}_i$ denotes the displacement components on $S^{\pm}$, then $u^+_i - u^-_i = b_i$. Equation (3.3) therefore

becomes

$$W = -\frac{1}{2}\int_S \sigma^D_{ij} m_j b_i dS - \int_S \sigma^B_{ij} m_j b_i dS + E^c, \tag{3.4}$$

where the integrals now take place over just one side of S, and are understood to be evaluated just up to a distance q from the dislocation. The two integrals written explicitly on the righthand side of (3.4) depend on the choice of the surface S. However, the term E^c also depends on S in such a way that the total energy W given in (3.4) is independent of the choice of S as it must be. Hence, as discussed by Bullough and Foreman[10], inclusion of the term E^c ensures uniqueness of the energy expression. If the background stress field σ^B is non-singular at the dislocation, which it usually is, then the contribution to E^c from the term involving σ^B may be neglected, being of the order of q, the core cutoff radius, which is assumed small.

We now move on to the definition of the driving force on the threading segment of a dislocation. Suppose that the threading arm translates with *fixed shape* through the stress field σ^B at speed v. The generalized force that drives the glide of the dislocation is denoted by G and is defined to be the force that is work-conjugate to the kinematic rate v with respect to the energy rate dW/dt. That is[22],

$$\frac{dW}{dt} = -Gv.$$

Then, as shown by Freund[22], differentiation of (3.4) yields the following approximate expression for the driving force:

$$G = -E^s + \int_T \sigma^B_{ij} m_j b_i d\eta, \tag{3.5}$$

where E^s is the energy per unit length of misfit segment being deposited (including the core-traction contribution), η is a coordinate defined in Fig. 3.1, and the integral takes place up the threading arm T. The quantity E^s may be approximated by the self-energy per unit length of an infinite straight misfit dislocation of the appropriate Burgers vector, given by (2.45).

Clearly, the force G is in general a function of the position $(x_1, x_2, x_3) = (x, y, z) = \mathbf{x}$ of the bottom of the threading arm. We shall therefore write $G = G(\mathbf{x})$. The interpretation to be placed on $G(\mathbf{x})$ is the following: if $G(x, y, z) > 0$ then motion of the threading arm at $\mathbf{x}$, depositing misfit dislocation at height y above the interface, causes a reduction in the energy of the system and hence tends to occur, with the converse holding if $G(\mathbf{x}) < 0$. With the above interpretation, insight into the strain relaxation process may be obtained by calculating $G(\mathbf{x})$ for typical background stress fields σ^B_{ij}.

3.2 Uniform Background Stress Field: Critical Layer Thickness

The simplest background stress field is that due solely to a uniform lattice mismatch e_{ij}^m, which has components given by (2.9). In this case the force G represents the force acting to move the threading arm of a dislocation so as to deposit the first misfit dislocation in an otherwise dislocation-free layer. The integral appearing in (3.5) is, for a uniform stress field, both trivial and independent of the shape of the threading arm of the dislocation. The force acting on a threading segment that is depositing misfit dislocation at the substrate-layer interface is in this case given by

$$G = -E^s - b_i m_j \Gamma_{ijkl} e_{kl}^m h \sec\alpha, \tag{3.6}$$

where

$$\Gamma_{ijkl} = c_{ijkl} - c_{ijr2}(\lambda^{-1})_{rs} c_{s2kl} \tag{3.7}$$

is the planar tensor of the moduli of elasticity for the direction (0,1,0) (see discussion following (2.9)). The sequential definition of critical thickness may now be given. For a given lattice mismatch e_{ij}^m, the critical layer thickness h_c is that thickness for which the force G, as given by (3.6), is zero. For $h < h_c$, G is negative, and so there is no force tending to drive the first dislocation across the layer. For $h > h_c$, G is positive, and so dislocation motion and strain relaxation are expected to commence. Thus h_c is viewed as a threshold thickness, which, for a given lattice mismatch, separates the regime in which dislocation-free growth is guarenteed from that in which threading-arm motion and consequent misfit dislocation formation occurs. This critical thickness condition was first annunciated by Matthews and Blakeslee[50].

By contrast, the energy minimization approach assumes that dislocations will be introduced so as to minimize the energy of the system. According to the equilibrium theory of van der Merwe[69], stable configurations are those in which the total energy is minimized. Thus, for a given layer thickness and lattice mismatch, the dislocation density is in general determined by the solution to the equation

$$\frac{dE}{d(1/p)} = 0, \tag{3.8}$$

where E is the total energy per unit area and $1/p$ is the linear dislocation density. For a stable configuration, the additional constraint must be satisfied that the stationary point be a local minimum, and it is usual to seek the configuration of lowest energy, that is the global minimum. It is possible that the lowest-energy configuration will not correspond to a stationary point of the total energy. For example, if the total energy of the system increases monotonically with

increasing dislocation density, then the dislocation-free layer will be the lowest-energy, and hence stable, configuration without necessarily corresponding to a stationary point of the total energy. Within this framework, the critical thickness has traditionally been defined to be, for a given lattice mismatch, the value of h for which (3.8) is satisfied in the limit as $p \to \infty$, corresponding to the introduction of a vanishing dislocation density, the idea being that a dislocation-free layer of the thickness so defined will be *just* stable.

In order to compare the sequential and energy minimization approaches to assessment of the critical layer thickness, it is convenient to move to a specific example, which is of great practical interest. Consider an $In_xGa_{1-x}As$ layer on a GaAs substrate. It is assumed that the epitaxial layer is grown in the (001) orientation, so that the x_2-axis coincides with the (001) direction for both materials. The materials considered are cubic, so the lattice mismatch takes the form

$$e_{ij}^m = f_m \delta_{ij}, \tag{3.9}$$

where f_m is the fractional lattice mismatch between the layer and substrate:

$$f_m = \frac{a_l - a_s}{a_s},$$

where a_l and a_s are the layer and substrate lattice constants respectively. It is assumed that the lattice constant of $In_xGa_{1-x}As$ follows Vegard's law. Taking the lattice constants of GaAs and InAs to be 5.65Å and 6.05Å respectively, it follows that the fractional lattice mismatch is related to the In fraction x by

$$f_m = 0.071x. \tag{3.10}$$

It is assumed that two orthogonal arrays, period p, of dislocations form between the epitaxial layer and its substrate. These are taken to be of the commonly observed '60°' type, which have a $\{111\}$ glide plane and a $\frac{1}{2}\langle 110\rangle$ Burgers vector. The dislocation line directions are the intersections of the $\{111\}$ glide planes with the (001) interface, namely $[110]$ and $[1\bar{1}0]$. The Burgers vectors of the $[1\bar{1}0]$ and $[110]$ dislocation lines are $\frac{1}{2}[\bar{1}01]$ and $\frac{1}{2}[011]$ respectively. Choosing the x_3-axis to coincide with a $[1\bar{1}0]$ dislocation line these Burgers vectors and their associated normals may be written

$$\mathbf{b} = (b_1, b_2, b_3) = b(-1/2, 1/\sqrt{2}, -1/2) \quad \text{with} \quad \mathbf{n} = (1, 0, 0)$$

and

$$\hat{\mathbf{b}} = (\hat{b}_1, \hat{b}_2, \hat{b}_3) = b(1/2, 1/\sqrt{2}, -1/2) \quad \text{with} \quad \hat{\mathbf{n}} = (0, 0, 1) \tag{3.11}$$

respectively. Here b is the magnitude of the Burgers vector; for GaAs, $b = 4$Å. In all calculations the core-cutoff radius is taken equal to b to ensure that a linear elastic model is not imposed on a region in which Hooke's law is grossly violated. In order to account for the non-linear core energy the core radius q is then reduced to $b/4$. The inclusion of non-linear core energies is discussed more fully in Section 5.5.

The stiffnesses and compliances are taken to be[61]

$$C_{1111} = 11.9 \times 10^{10}\text{Pa}, \quad C_{1122} = 5.40 \times 10^{10}\text{Pa}, \quad C_{1212} = 5.95 \times 10^{10}\text{Pa},$$
$$S_{1111} = 0.117 \times 10^{-10}\text{Pa}^{-1}, \quad S_{1122} = -0.0365 \times 10^{-10}\text{Pa}^{-1}, \quad S_{1212} = 0.168 \times 10^{-10}\text{Pa}^{-1},$$

for GaAs and

$$C_{1111} = 8.39 \times 10^{10}\text{Pa}, \quad C_{1122} = 4.50 \times 10^{10}\text{Pa}, \quad C_{1212} = 3.99 \times 10^{10}\text{Pa},$$
$$S_{1111} = 0.194 \times 10^{-10}\text{Pa}^{-1}, \quad S_{1122} = -0.680 \times 10^{-10}\text{Pa}^{-1}, \quad S_{1212} = 0.253 \times 10^{-10}\text{Pa}^{-1},$$

for InAs. To obtain stiffnesses for the alloy $\text{In}_x\text{Ga}_{1-x}\text{As}$, both the compliances and stiffnesses were first interpolated between their values for GaAs and InAs. The compliances thus obtained were then inverted to yield a second estimate for the stiffnesses. Taking the mean of these two estimates for each stiffness gave the final value used for the alloy. The constants given above have been denoted by upper-case letters to emphasize that they are referred to the cubic axes of the crystals. The elastic constants must be transformed so that they are referred to the coordinate system illustrated in Fig. 2.1. In the case studied here, this coordinate system has the x_1, x_2, and x_3-axes running in directions [110], [001], and [1$\bar{1}$0] respectively. Referred to these axes the non-zero stiffnesses are

$$\begin{aligned}
c_{1111} &= c_{3333} = \frac{1}{2}\left(C_{1111} + C_{1122} + 2C_{1212}\right), \\
c_{2222} &= C_{1111}, \\
c_{1122} &= c_{2233} = C_{1122}, \\
c_{1133} &= \frac{1}{2}\left(C_{1111} + C_{1122} - 2C_{1212}\right), \\
c_{1212} &= c_{2323} = C_{1212}, \\
c_{1313} &= \frac{1}{2}\left(C_{1111} - C_{1122}\right).
\end{aligned}$$

The system has the orthorhombic symmetry described in Section 2.7, and so the γ_r are simply the roots of (2.53) that have positive imaginary part.

In a cubic material, the [110] and [1$\bar{1}$0] directions (along which the dislocations lie) are elastically equivalent, and so $\hat{E}^{fa} = E^{fa}$ (see discussion after (2.47)). Therefore, the total energy per unit area may be written, using (2.46), as

$$E = \frac{2}{p}E^{fa} + \frac{1}{2}h\Gamma_{ijkl}e^r_{kl}e^r_{ij}, \tag{3.12}$$

since, as discussed after (2.16) and (2.47),

$$\bar{\sigma}_{ij} = -\Gamma_{ijkl}e^r_{kl}.$$

The relaxed lattice mismatch, e^r_{ij}, is given by (2.47) with $\mathbf{b}$, $\hat{\mathbf{b}}$, $\mathbf{n}$, and $\hat{\mathbf{n}}$ given by (3.11). The Burgers vectors of the two arrays have been chosen so that the long-range shear strains due to the screw components of the two arrays combine to produce a small rigid rotation of the layer. Another way to avoid long-range shear is to alternate the screw components of dislocations within each array. The energy for such a case may be obtained from Section 2.6.

Now, as may be seen explicitly from (3.7), Γ_{ijkl} possesses all of the usual symmetries of the elastic constants:

$$\Gamma_{ijkl} = \Gamma_{klij} = \Gamma_{lkij}.$$

Therefore,

$$\frac{1}{2}\frac{d}{d(1/p)}\Gamma_{ijkl}e^r_{kl}e^r_{ij} = \Gamma_{ijkl}e^r_{kl}\frac{d}{d(1/p)}e^r_{ij} = \Gamma_{ijkl}e^r_{kl}(b_in_j + \hat{b}_i\hat{n}_j),$$

where the explicit form of e^r_{ij}, from (2.47), has been used. Therefore, differentiation of (3.12) yields

$$\frac{dE}{d(1/p)} = 2E^{fa} + \frac{2}{p}\frac{dE^{fa}}{d(1/p)} + h\Gamma_{ijkl}e^r_{kl}(b_in_j + \hat{b}_i\hat{n}_j). \tag{3.13}$$

It may be seen explicitly from (2.44) that

$$\frac{1}{p}\frac{dE^{fa}}{d(1/p)} \to 0 \quad \text{as } p \to \infty.$$

Therefore, using (3.13), (3.8) yields, on taking the limit as $p \to \infty$,

$$2E^{fa} + hf_m\Gamma_{ijkk}(b_in_j + \hat{b}_i\hat{n}_j) = 0, \tag{3.14}$$

where the particular form, (3.9), for e^m_{ij} has been used. Since the [110] and [1$\bar{1}$0] directions are, elastically, equivalent in a cubic material such as InGaAs, $\Gamma_{11kk} = \Gamma_{33kk}$ for the system under

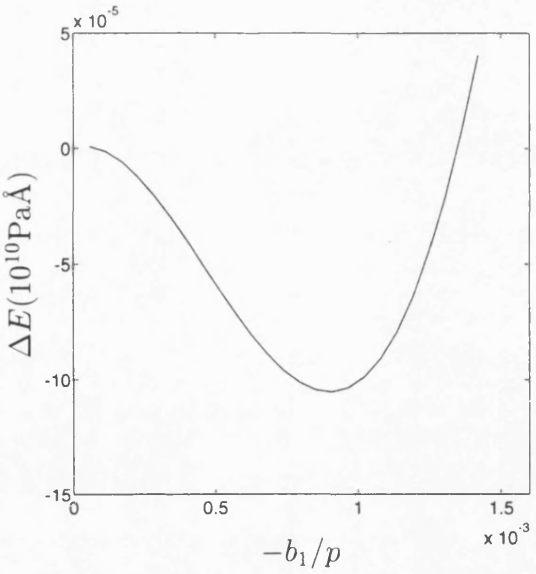

Figure 3.2: The local structure of the stationary point of the energy variation as $p \to \infty$ at the critical thickness h_c in an $In_{0.1}Ga_{0.9}As$ layer on GaAs(001), as demonstrated by the energy change ΔE per unit area on introduction of two orthogonal arrays of 60° dislocations, plotted as a function of $1/p$.

consideration. Therefore, (3.14) reduces to

$$E^{fa} + hf_m\Gamma_{ijkk}b_i n_j = 0. \tag{3.15}$$

Now $\Gamma_{i2kl} = 0$, and $n_1 = m_1 \sec\alpha$. Therefore, (3.15) is identical with the condition

$$G = 0, \tag{3.16}$$

with G given by (3.6). Hence, for a system with the symmetry considered here, equating to zero the force of (3.6) yields a critical thickness identical with that obtained by solving (3.8) in the limit as $p \to \infty$. This equivalence was, to the knowledge of the author, first noted by Willis *et al.*[74]. However, as we shall now demonstrate, a layer of this critical thickness is energetically *unstable.*

In Fig. 3.2 the energy change, from that of a dislocation-free layer, is plotted as a function of $1/p$ for large p. A layer of $In_{0.1}Ga_{0.9}As$ on GaAs is considered, of thickness equal to the critical thickness, as defined by (3.15), or equivalently (3.16). The stationary point of the energy as $p \to \infty$ is found to be a *maximum*, not a minimum as was previously thought. This effect was first noted for the case of isotropy by Jain *et al.*[38] and has been examined in detail for

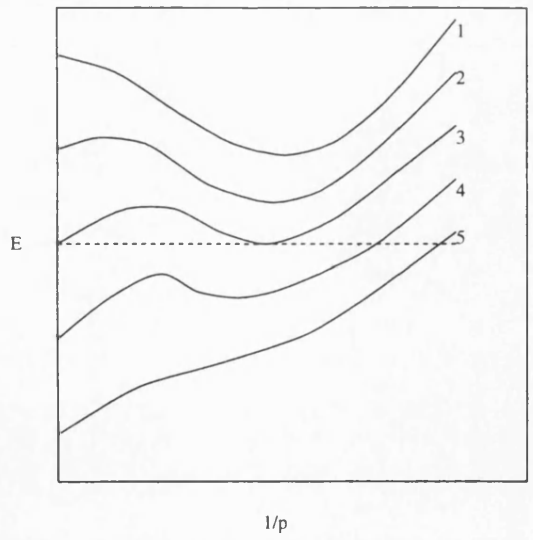

Figure 3.3: Schematic representation of energy variation with dislocation density: curve 3 gives the energy variation at the correct energy-minimization critical thickness; the layer thickness *decreases* from curve 1 through to curve 5.

this case by Gosling *et al.*[26]. In the isotropic case, the effect was found to be dependent upon the value of Poisson's ratio, disappearing for $\nu > 1/3$ when 60° dislocations are considered. It is therefore interesting to observe that the effect persists (and is in fact enhanced[27]) when the full anisotropic solution is used. This behaviour does not affect the sequential definition of the critical thickness, but it does affect the energy minimization definition. At the thickness obtained by solving (3.8) in the limit as $p \to \infty$ the lowest energy configuration corresponds to a finite dislocation spacing, as illustrated in Fig. 3.2, so if the energy is always kept at a minimum then the strain relaxation process will already have commenced. This thickness is therefore not the correct energy-minimization critical thickness. If the energy minimization approach is followed then the energy variation with dislocation density in a layer of the critical thickness must be as illustrated in curve 3 of Fig. 3.3. At this thickness the energy of the dislocation-free layer is precisely the same as that of the lowest-energy configuration for a finite dislocation spacing. For smaller thicknesses the dislocation-free layer the is global lowest-energy configuration (see curves 4 and 5 of Fig. 3.3). For larger thicknesses there will be a lower-energy state corresponding to a finite dislocation density, and so precisely this density of dislocations should be introduced (see curves 1 and 2 of Fig. 3.3).

The correct energy-minimization critical thickness defined above is always less than that obtained by solving (3.15) and (3.16), but only by a few ångströms, which, for practical purposes, is a negligible difference. However, it does illustrate a fundamental difference between the energy

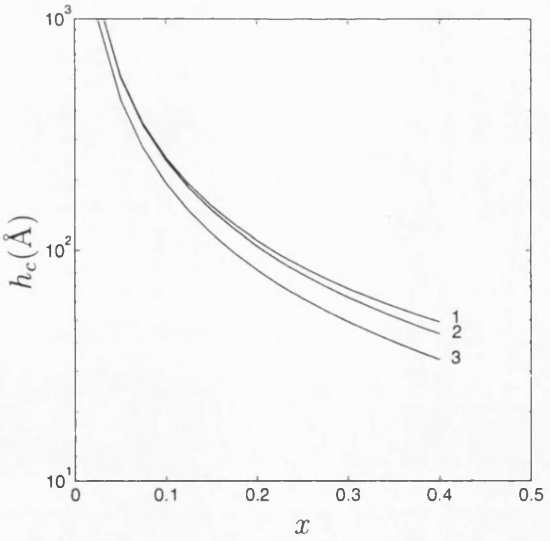

Figure 3.4: Theoretical critical thickness of an $In_xGa_{1-x}As$ layer grown on GaAs(001) plotted as a function of the In fraction x: curve 1, h_c^{an} obtained using the anisotropic solution including the correction term to account for the inhomogeneity of the body; curve 2, anisotropic homogeneous body value h_c^{an0}, assuming that the elastic constants are those of GaAs throughout; curve 3, h_c^{is0}, obtained using the isotropic homogeneous body approximation.

minimization and sequential approaches. If a small change in configuration causes an increase in energy, then the sequential approach will not allow a change in that direction. However, the energy minimization approach allows a lowest-energy configuration to be 'seen' over an energy barrier.

The theoretical critical thickness of an $In_xGa_{1-x}As$ layer on a GaAs(001) substrate, as obtained from (3.16) is plotted in Fig. 3.4 as a function of In fraction x. It has been assumed that relaxation takes place via the introduction of 60° dislocations. Curve 1 gives the critical thickness h_c^{an}; in calculating this quantity the first order correction to allow for the inhomogeneity of the body has been incorporated. Curve 2 gives the critical thickness h_c^{an0} obtained from the anisotropic, homogeneous body solution, assuming that the elastic constants are those of GaAs throughout, and curve 3 gives h_c^{is0}, which is obtained from the isotropic, homogeneous body approximation[74, 26]. Incorporation of the anisotropic nature of the material results in a larger predicted critical thickness than is obtained from the isotropic approximation, and this value is increased further on allowing for the difference between the elastic properties of the layer and substrate. This is emphasized in Fig.3.5 where h_c^{an} and h_c^{an0}, normalized by the isotropic critical thickness h_c^{is0}, are plotted as a function of x. Use of the anisotropic, homogeneous body solution results in an increase of 25%-30%, compared with the isotropic approximation, in

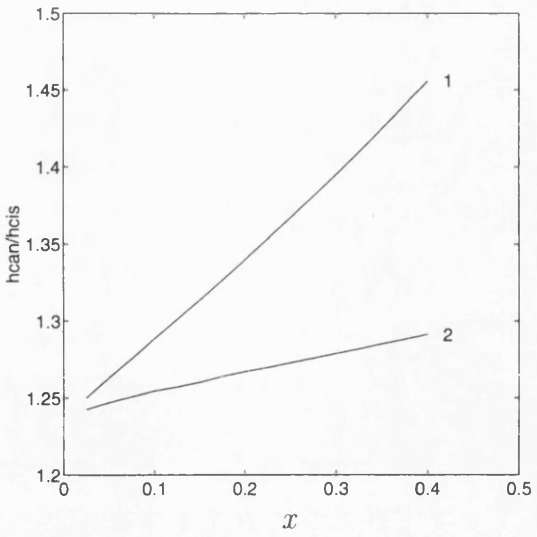

Figure 3.5: Plot of the ratios h_c^{an}/h_c^{is0} (curve 1) and h_c/h_c^{is0} (curve 2) as functions of the In fraction x; illustrated is the increase in predicted critical thickness, compared with that obtained using the isotropic homogeneous body approximation, that results from incorporation of anisotropy (curve 2) and both anisotropy and inhomogeneity (curve 1).

the predicted critical thickness. This increase is as much as 45% when the correction term to account for the inhomogeneity of the body is also incorporated.

When the elastic constants of the substrate differ from those of the layer, the energy of a dislocation might be reduced if it is displaced slightly into the softer material. This would affect, for example, the critical thickness of the layer. This effect has been neglected here, because only a small difference between the elastic constants of the layer and substrate was considered theoretically. Moreover, for the particular material system studied here, namely $In_xGa_{1-x}As/GaAs$, in which the strained layer is the softer material, the force due to the mismatch strain, which drives the dislocation towards the interface, militates against displacement of the dislocation into the softer strained layer.

3.3 Partially Relaxed Layers and Stable Configurations

We now turn our attention to consideration of the ongoing process of strain relaxation in layers of thickness greater than the critical value discussed in the previous section. It is obvious that as dislocations are introduced into a layer, the strain in the layer is relaxed, and hence the driving force for propagation of further dislocations is reduced. Eventually the driving force will be insufficient for the relaxation of strain to continue. A configuration that resists the

introduction or removal of dislocations is said to be stable. The purpose of this section is to make this definition mathematically more precise. The number of possible configurations is so enormous that it is not feasible to cover them comprehensively. In this section we therefore restrict ourselves to a subset of the possible configurations, and at the end of the chapter we shall consider how the situation may differ for configurations outside this subset. As in the previous section we shall consider an $In_xGa_{1-x}As/GaAs(001)$ strained layer system with two orthogonal arrays, A and $\hat{A}$, of misfit dislocations of period p at the interface. Dislocations in the arrays A and $\hat{A}$ have Burgers vectors $\mathbf{b}$ and $\hat{\mathbf{b}}$, respectively, with associated normals $\mathbf{n}$ and $\hat{\mathbf{n}}$, as given by (3.11). We shall ask for what values of p the above configuration is stable, for given In fraction x and layer thickness h. Different answers to this question will follow from the energy-minimization and sequential approaches.

According to the energy minimization approach, the array configuration described above is stable if and only if the spacing p is a solution of (3.8), for the given lattice mismatch and layer thickness. The derivative to be set to zero in (3.8) may be obtained from Chapter 2, via (3.13). Although attractive because of its simplicity, this approach has weaknesses, most notably in that it fails to incorporate any features of the mechanism by which strain relaxation occurs. The sequential approach allows a more mechanisitic view to be taken, based on considering the extension or contraction of a segment of misfit dislocation via the motion of its associated threading segment. On this view the array configuration considered may change either through the removal of a misfit dislocation already in the array or through the introduction of a new misfit dislocation between two members of the array. The configuration is stable if there is no driving force for either of these two changes to occur. By combining the principles of Section 3.1 with the solutions of Chapter 2 we shall now calculate the driving force for these two occurences.

The glide plane of a dislocation of the commonly observed 60° type is inclined to the (001) interface, and so adjustment of the position of such a dislocation in the interface can only take place by climb, which requires the diffusion of interstitials and vacancies to the dislocation. The timescale for adjustment of dislocation positions in the interface is therefore several orders of magnitude greater than that for glide of the threading arm of a dislocation[41]. It is therefore reasonable to suppose that misfit dislocations already present in the layer remain fixed, unable to adjust their positions in the interface, as a further dislocation propagates across the layer. Therefore the forces we require are those acting on the threading arm of a dislocation moving in fixed background stress fields, which forces may be calculated using (3.5).

Proceeding to details, we require the force G_1 acting to drive a dislocation between two members of one of the two orthogonal arrays, and we require the force G_2 acting to move the

threading arm of a dislocation so as to contract and remove its associated misfit segment in one of the two arrays, thereby creating a vacancy in the periodic structure. Let the origin of coordinates lie at the intersection of two orthogonal dislocations. Let $\mathbf{x}=(x_1, x_2, x_3)$ be coordinates referred to the coordinate system of array A so that the dislocations of array A lie parallel to the x_3-axis, and the associated normal $\mathbf{n}$ is parallel to $0x_1$. The dislocations of array $\hat{A}$ then lie parallel to the x_1-axis, and $\hat{\mathbf{n}}$ runs in the direction of $0x_3$. We consider the addition or removal of a dislocation parallel to the x_3-axis. Of course, because of the array perpendicular to the propagating dislocation, the force on its threading arm will vary with x_3. Moreover, as a result of this, the force on the threading arm will depend upon its shape. For simplicity, we assume that the threading arm minimizes its length within the glide plane. In other words, the threading arm is assumed to be straight, to lie in the glide plane, and to be perpendicular to the misfit segment that it is depositing, as illustrated in Fig. 3.1. The forces G_1 and G_2 are now calculated by inserting appropriate background stress fields σ^B in (3.5).

First consider G_1. The background stress field in this case is that due to two orthogonal arrays of dislocations and the lattice mismatch. These stress fields are given in Sections 2.2 and 2.3. For simplicity we consider a misfit dislocation being deposited via threading motion exactly half-way between two members of the array A (i.e. at position $x_1 = p/2$), which lie parallel to the x_3-axis. Let the Burgers vector of the propagating dislocation be $\tilde{\mathbf{b}}$. Denote by σ_{ij}^{fa} and $\hat{\sigma}_{ij}^{fa}$ the fluctuating parts of the stresses due to arrays A and $\hat{A}$, respectively. Then (3.5) gives

$$G_1(y,z) = -E^s(h-y) - \tilde{b}_i m_j \Gamma_{ijkl} e_{kl}^r h \sec\alpha + \int_T \sigma_{ij}^{fa} m_j \tilde{b}_i d\eta + \int_T \hat{\sigma}_{ij}^{fa} m_j \tilde{b}_i d\eta, \qquad (3.17)$$

where $E^s(h-y)$ is the energy per unit length of misfit dislocation at height y above the interface (i.e. at distance $h-y$ from the free surface). The base of the threading arm T is taken to be half way between two members of the array A; that is, at $\mathbf{x}=(p/2, y, z)$. Now the stress field σ_{ij}^{fa} does not vary with x_3, so by the equilibrium equations

$$\sigma_{i\alpha,\alpha}^{fa} = 0.$$

Elementary use of the divergence theorem then reveals that

$$\int_T \sigma_{ij}^{fa} m_j \tilde{b}_i d\eta = \int_y^h \sigma_{i1}^{fa} \tilde{b}_i dx_2,$$

where the righthand side is evaluated at $x_1 = p/2$. Now, recalling from Chapter 2 the notation

$$\sigma_{ij}^{fa} = t_i^{(j)},$$

we may write

$$\begin{aligned}\int_y^h \sigma_{i1}^{fa}\tilde{b}_i dx_2 &= \int_y^h \tilde{\mathbf{b}}^{\mathrm{T}}\mathbf{t}^{(1)}dx_2 \\ &= \frac{1}{p}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \tilde{\mathbf{b}}^{\mathrm{T}}\left[\int_y^h \hat{\mathbf{t}}^{(1)}\left(\frac{2\pi k}{p}\right)dx_2\right]\mathrm{e}^{-2\pi \mathrm{i} k x_1/p}.\end{aligned}$$

The term in square brackets is known from (2.38) and (2.39) by putting $q = y$ there. Evaluating at $x_1 = p/2$ and using the fact that $\hat{\mathbf{u}}(-\xi) = \overline{\hat{\mathbf{u}}(\xi)}$ we obtain

$$\int_y^h \sigma_{i1}^{fa}\tilde{b}_i dx_2 = -\frac{4}{p}\mathrm{Re}\sum_{k=1}^{\infty}\tilde{\mathbf{b}}^{\mathrm{T}}(-1)^k\mathbf{f}^0\left(\frac{2\pi k}{p}\right), \tag{3.18}$$

where q is replaced by y in $\mathbf{f}^0$. Completion of the summation as required by (3.18) yields

$$\begin{aligned}\int_T \sigma_{ij}^{fa}m_j\tilde{b}_i d\eta &= -\mathrm{Im}\frac{\tilde{\mathbf{b}}^{\mathrm{T}}}{\pi}\left\{\sum_{r,s}\frac{\mathbf{C}_r^{(1)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s(\mathbf{C}_s^{(2)})^{\mathrm{T}}}\ln\left[\frac{1+\mathrm{e}^{-2\pi\mathrm{i}[(\overline{\gamma_s}-\gamma_r)h+\gamma_r y]/p}}{1+\mathrm{e}^{-2\pi\mathrm{i}\overline{\gamma_s}h/p}}\right]\right. \\ &\left. + \sum_r \frac{\overline{\mathbf{C}_r^{(1)}\mathbf{N}_r\mathbf{R}_r}}{\overline{\gamma_r}}\ln\left(\frac{1+\mathrm{e}^{-2\pi\mathrm{i}\overline{\gamma_r}h/p}}{1+\mathrm{e}^{-2\pi\mathrm{i}\overline{\gamma_r}y/p}}\right)\right\}\mathbf{b}.\end{aligned} \tag{3.19}$$

Now the stresss field $\hat{\sigma}_{ij}^{fa}$ has no variation with respect to x_1, so

$$\int_T \hat{\sigma}_{ij}^{fa}m_j\tilde{b}_i d\eta = \sec\alpha\int_y^h \hat{\sigma}_{ij}^{fa}m_j\tilde{b}_i dx_2. \tag{3.20}$$

Let $\mathbf{x}' = (x_1', x_2', x_3')$ be coordinates referred to the coordinate system of array $\hat{A}$, which has the x_3'-axis running parallel to the dislocations in this array, and the x_1'-axis running in the direction of $\hat{\mathbf{n}}$. Let $\mathbf{M}$ be the matrix that transforms coordinates in the system of $\hat{A}$ to those in the system of A so that

$$\mathbf{x} = \mathbf{M}\mathbf{x}'.$$

Then

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Now $\hat{\mathbf{b}}' = \mathbf{M}^{\mathrm{T}}\hat{\mathbf{b}} = \mathbf{b}$ and $\hat{\mathbf{n}}' = \mathbf{M}^{\mathrm{T}}\hat{\mathbf{n}} = \mathbf{n}$ so in the coordinate system of $\hat{A}$ we have

$$(\hat{\sigma}^{fa})'_{ij} = \sigma^{fa}_{ij}.$$

Therefore, in the coordinate system of A,

$$\hat{\sigma}^{fa}_{ij}(\mathbf{x}) = M_{ip}M_{jq}\sigma^{fa}_{pq}(\mathbf{M}^{\mathrm{T}}\mathbf{x}).$$

Therefore

$$\begin{aligned}\int_y^h \hat{\sigma}^{fa}_{ij}(\mathbf{x})m_j\tilde{b}_i dx_2 &= \int_y^h \sigma^{fa}_{pq}(\mathbf{M}^{\mathrm{T}}\mathbf{x})M_{jq}m_jM_{ip}\tilde{b}_i dx_2 \\ &= \int_y^h t^{(q)}_p(\mathbf{M}^{\mathrm{T}}\mathbf{x})M_{jq}m_jM_{ip}\tilde{b}_i dx_2. \end{aligned} \tag{3.21}$$

Now, $\mathbf{M}^{\mathrm{T}}\mathbf{m} = (0, m_2, -m_1)$, so (3.21) may be written

$$\int_y^h \hat{\sigma}^{fa}_{ij}(\mathbf{x})m_j\tilde{b}_i dx_2 = \int_y^h \tilde{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\left[\mathbf{t}^{(2)}(\mathbf{M}^{\mathrm{T}}\mathbf{x})m_2 - \mathbf{t}^{(3)}(\mathbf{M}^{\mathrm{T}}\mathbf{x})m_1\right] dx_2. \tag{3.22}$$

The Fourier coefficient of $\mathbf{t}^{(k)}$ is known from (2.27). Integration yields:

$$\int_y^h \tilde{\mathbf{t}}^{(k)}dx_2 = \sum_r \left[\frac{\mathbf{C}^{(k)}_r\mathbf{N}_r\mathbf{a}^0_r}{\gamma_r}(\mathrm{e}^{-\mathrm{i}\xi\gamma_r h} - \mathrm{e}^{-\mathrm{i}\xi\gamma_r y}) + \frac{\overline{\mathbf{C}^{(k)}_r\mathbf{N}_r\mathbf{b}^0_r}}{\overline{\gamma_r}}(\mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_r}h} - \mathrm{e}^{-\mathrm{i}\xi\overline{\gamma_r}y})\right],$$

where $\mathbf{a}^0_r$ and $\mathbf{b}^0_r$ are given by (2.31) and (2.32). Now $\mathbf{M}^{\mathrm{T}}\mathbf{x} = (x_3, x_2, -x_1)$, and

$$\int_y^h \mathbf{t}^{(k)}(x_3, x_2)dx_2 = \frac{2}{p}\mathrm{Re}\sum_{\substack{k=1\\ \xi=2\pi k/p}}^{\infty} \mathrm{e}^{-\mathrm{i}\xi x_3}\int_y^h \tilde{\mathbf{t}}^{(k)}(\xi, x_2)dx_2.$$

Inserting expressions from (2.31) and (2.32) for $\mathbf{a}^0_r$ and $\mathbf{b}^0_r$ and summing the series using (2.43) yields:

$$\begin{aligned}\int_y^h \mathbf{t}^{(k)}(x_3, x_2)dx_2 &= -\mathrm{Im}\frac{1}{\pi}\left\{\sum_{r,s}\frac{\mathbf{C}^{(k)}_r\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s(\mathbf{C}^{(2)}_s)^{\mathrm{T}}}\ln\left[\frac{1-\mathrm{e}^{-2\pi\mathrm{i}[(\overline{\gamma_s}-\gamma_r)h+\gamma_r y+x_3]/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_s}h+x_3)/p}}\right]\right. \\ &\quad \left. + \sum_r \frac{\overline{\mathbf{C}^{(k)}_r\mathbf{N}_r\mathbf{R}_r}}{\overline{\gamma_r}}\ln\left[\frac{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_r}h+x_3)/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_r}y+x_3)/p}}\right]\right\}\mathbf{b}.\end{aligned}$$

The matrices $\mathbf{N}$, $\mathbf{C}^{(k)}$, $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ are defined in (2.22), (2.25), (2.34), and (2.40).

Equations (3.20) and (3.22) now give, in the coordinate system of A,

$$\begin{aligned}\int_T \hat{\sigma}_{ij}^{fa} m_j \tilde{b}_i d\eta \;=\; & -\mathrm{Im}\frac{\tilde{\mathbf{b}}^{\mathrm{T}}\mathbf{M}}{\pi}\left\{\left[\sum_{r,s}\frac{\mathbf{C}_r^{(2)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s(\mathbf{C}_s^{(2)})^{\mathrm{T}}}\ln\left(\frac{1-\mathrm{e}^{-2\pi\mathrm{i}[(\overline{\gamma_s}-\gamma_r)h+\gamma_r y+z]/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_s}h+z)/p}}\right)\right.\right. \\ & \left.+\sum_r\frac{\overline{\mathbf{C}_r^{(2)}\mathbf{N}_r\mathbf{R}_r}}{\overline{\gamma_r}}\ln\left(\frac{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_r}h+z)/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_r}y+z)/p}}\right)\right]m_2\sec\alpha \\ & -\left[\sum_{r,s}\frac{\mathbf{C}_r^{(3)}\mathbf{N}_r}{\gamma_r}\mathbf{P}^{-1}\overline{\mathbf{Q}_s(\mathbf{C}_s^{(2)})^{\mathrm{T}}}\ln\left(\frac{1-\mathrm{e}^{-2\pi\mathrm{i}[(\overline{\gamma_s}-\gamma_r)h+\gamma_r y+z]/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_s}h+z)/p}}\right)\right. \\ & \left.\left.+\sum_r\frac{\overline{\mathbf{C}_r^{(3)}\mathbf{N}_r\mathbf{R}_r}}{\overline{\gamma_r}}\ln\left(\frac{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_r}h+z)/p}}{1-\mathrm{e}^{-2\pi\mathrm{i}(\overline{\gamma_r}y+z)/p}}\right)\right]m_1\sec\alpha\right\}\mathbf{b}. \qquad (3.23)\end{aligned}$$

Equations (3.19) and (3.23) complete the definition of the force G_1 given in (3.17).

We now want the force G_2 acting to drive a dislocation out of the array A creating a vacancy. Define this force to be positive if a dislocation in the array tends to contract. Note that this is just minus the force acting on a dislocation of Burgers vector $\mathbf{b}$ so as to move it into a vacancy in the periodic structure of array A. This latter force is obtained from (3.5) by inserting a background stress field consisting of contributions from periodic arrays perpendicular and parallel to the propagating dislocation *less* the stresses due to a single dislocation at the position of the vacancy, When considering the contraction of a misfit dislocation lying at the interface, we are interested only in the force for $y = 0$. The above considerations lead to the conclusion that

$$G_2(z) = E^s(h) + b_i m_j \Gamma_{ijkl} e_{kl}^r h\sec\alpha - \int_T(\sigma_{ij}^{fa} - \sigma_{ij}^s)m_j b_i dx_2' - \int_T \hat{\sigma}_{ij}^{fa} m_j b_i dx_2'. \qquad (3.24)$$

The final integral appearing on the righthand side of (3.24) may be obtained from (3.23) by taking the case $y = 0$ there; only the penultimate integral remains to be evaluated. Of interest is the force acting to remove a dislocation from the array A, so the threading arm T runs from a position $x_1 = np$ at the interface to the free surface. Therefore σ_{ij}^{fa} and σ_{ij}^s are both singular at the base of the threading arm. However, the difference $\sigma_{ij}^{fa} - \sigma_{ij}^s$ is finite there, and considerations similar to those following (3.17) reveal that

$$-\int_T(\sigma_{ij}^{fa} - \sigma_{ij}^s)m_j b_i d\eta = -\int_0^h(\sigma_{i1}^{fa} - \sigma_{i1}^s)b_i dx_2, \qquad (3.25)$$

where the integral on the righthand side is to be evaluated at $x_1 = 0$. Now

$$-\int_0^h (\sigma_{i1}^{fa} - \sigma_{i1}^{s}) b_i dx_2 = 2 \lim_{q \to 0} \left\{ -\frac{1}{2} \int_q^h \sigma_{i1}^{fa} b_i dx_2 + \lim_{p \to \infty} \frac{1}{2} \int_q^h \sigma_{i1} b_i dx_2 \right\}.$$

But from (2.36) we see that

$$-\frac{1}{2} \int_q^h \sigma_{i1}^{fa} b_i dx_2 = E^{fa} - E^c;$$

moreover,

$$\lim_{p \to \infty} (E^{fa} - E^c) = E^s - E^c.$$

Therefore, (3.25) may be written

$$-\int_T (\sigma_{ij}^{fa} - \sigma_{ij}^{s}) m_j b_i d\eta = 2 \lim_{q \to 0} (E^{fa} - E^s),$$

or, using (2.44) and (2.45),

$$\begin{aligned} -\int_T (\sigma_{ij}^{fa} - \sigma_{ij}^{s}) m_j b_i d\eta \;=\; & \operatorname{Im} \frac{\mathbf{b}^{\mathrm{T}}}{\pi} \left\{ \sum_{r,s} \frac{\mathbf{C}_r^{(1)} \mathbf{N}_r}{\gamma_r} \mathbf{P}^{-1} \overline{\mathbf{Q}_s (\mathbf{C}_s^{(2)})^{\mathrm{T}}} \left[\ln \left(\frac{1 - \mathrm{e}^{-2\pi \mathrm{i} (\overline{\gamma_s} - \gamma_r) h/p}}{1 - \mathrm{e}^{-2\pi \mathrm{i} \overline{\gamma_s} h/p}} \right) \right. \right. \\ & \left. - \ln \left(\frac{\overline{\gamma_s} - \gamma_r}{\overline{\gamma_s}} \right) \right] \\ & \left. + \sum_r \frac{\overline{\mathbf{C}_r^{(1)} \mathbf{N}_r \mathbf{R}_r}}{\overline{\gamma_r}} \ln \left(\frac{p(1 - \mathrm{e}^{-2\pi \mathrm{i} \overline{\gamma_r} h/p})}{2\pi \mathrm{i} \overline{\gamma_r} h} \right) \right\} \mathbf{b}. \end{aligned} \tag{3.26}$$

Equation (3.26) completes the definition of G_2 given in (3.24).

Before proceeding to the precise mathematical definition of stability for a periodic array, some observations concerning interpretation of the forces G_1 and G_2 are of relevance. First consider the propagation of a dislocation depositing misfit dislocation half-way between two members of array A. As a rule, an exception to which is discussed next, the threading arm will deposit misfit dislocation at the interface (where its strain-relieving effect is maximized) so the force of relevance is $G_1(0, z)$. However, as the propagating dislocation approaches a perpendicular misfit dislocation in the array $\hat{A}$, the driving foce $G_1(0, z)$ may be reduced to zero for some value of z, because of the interaction with the perpendicular misfit dislocation. The propagating dislocation can only proceed past the obstacle by depositing misfit dislocation at some height y above the interface for which the driving force is positive. This process is illustrated in Fig. 3.6, where the threading arm stops depositing misfit dislocation at the interface when $z = z^\star$. In the case illustrated, the force for deposition at height $y = y^\star$ is always positive, and the propagating dislocation is able to pass the perpendicular misfit dislocation. If, at some z, the driving force

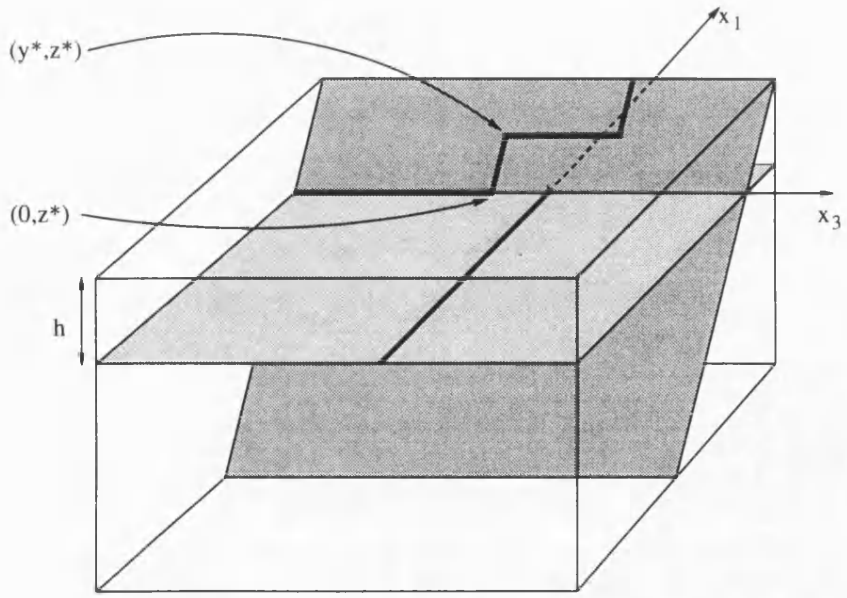

Figure 3.6: Schematic illustration of the threading arm of a dislocation depositing misfit dislocation at height y^* above the interface in order to pass a perpendicular misfit dislocation.

is zero or negative for deposition of misfit dislocation at all heights $y < h$ above the interface, then motion of the propagating dislocation is halted.

Now, as the dislocation begins to deposit misfit segment at height y^* above the interface, Freund's derivation of the driving force[22] is not strictly valid. This is because that derivation relies on the stress field being essentially two-dimensional above *some* point on the misfit segment being deposited. This requirement is certainly not fulfilled when the misfit segment at height y^* above the interface is very short. However, in spite of this, we asume that (3.5) may be used to a reasonable approximation. Also note that if the propagating dislocation initially deposits misfit dislocation in the interface at position $x_1 = p/2$, exactly half-way between two members of the array, then when depositing misfit dislocation at height y above the interface, its position will be off-centre, at position $x_1 = p/2 \pm y \tan\alpha$ because of the inclination of the glide plane to the vertical. Since the glide plane in the system coonsidered is inclined at not too severe an angle to the vertical ($\tan\alpha = 1/\sqrt{2}$), and since there are other uncertainties in the model, we ignore this detail, and take the force on the dislocation always to be that for central deposition of misfit dislocation.

Bearing in mind the above discussion, we are now in a position to give the sequential definition of array stability. Consider a pair of orthogonal arrays of dislocations of period p in a strained-layer system of given lattice mismatch and layer thickness, as described at the beginning of this section. This configuration is stable provided that

Case	$\tilde{\mathbf{b}}/b$
(i)	$(-1/2, 1/\sqrt{2}, 1/2)$
(ii)	$(-1/2, 1/\sqrt{2}, -1/2)$
(iii)	$(-1/2, -1/\sqrt{2}, 1/2)$
(iv)	$(-1/2, -1/\sqrt{2}, -1/2)$

Table 3.1: Possibilities for the Burgers vector $\tilde{\mathbf{b}}$ of the propagating dislocation.

(a) there is a fixed z such that $G_1(y,z) \leq 0$ for all $y < h$;

(b) $G_2 \leq 0$ for some z.

Condition (a) ensures that a further dislocation cannot be deposited via threading motion between two members of the array, and condition (b) ensures that no misfit dislocation already in the array will contract under its own line tension and disappear. Let p_a be the *largest* value of dislocation spacing p for which condition (a) is satisfied, and let p_b be the *smallest* value of p for which condition (b) is satisfied. Then it is found that the periodic array configuration is stable for $p_b < p < p_a$. Moreover, $p_b < p_m < p_a$, where p_m is the spacing obtained by minimizing the energy of the system, i.e. by solving (3.8). Therefore, when the sequential approach is adopted, the concept of a unique stable configuration does not exist. Instead there is a range of spacings p for which the periodic array configuration is stable, this range containing the spacing p_m of the unique energy-minimizing configuration.

Of primary interest is the spacing p_a defined by condition (a). This spacing gives, for fixed lattice mismatch and layer thickness, the dislocation density at which introduction of further dislocations will first be resisted. We now calculate p_a for the $In_xGa_{1-x}As/GaAs(001)$ system, assuming the introduction of 60° dislocations. With $\mathbf{b}$ and $\hat{\mathbf{b}}$ as given by (3.11), there are four strain-relieving choices for $\tilde{\mathbf{b}}$. These are listed in Table 3.1. It is found that for case (i) the interaction between the propagating dislocation and those perpendicular to it is repulsive and can give rise to blocking of the dislocation. For the other three cases the interaction is predominantly attractive and so no such blocking is likely to occur. Of course, our assumption that all dislocations in the array $\hat{A}$ have the same Burgers vector $\hat{\mathbf{b}}$ is an idealization. One would expect the Burgers vectors of members of $\hat{A}$ to be fairly evenly distributed amongst the available strain relieving values. Symmetry arguments show that the four pairings $(\hat{\mathbf{b}},\tilde{\mathbf{b}})$, with $\hat{\mathbf{b}}$ given by (3.11) and $\tilde{\mathbf{b}}$ given by one of the cases in Table 3.1, represent the four independent interactions between perpendicular dislocations. Therefore, a propagating dislocation of any Burgers vector $\tilde{\mathbf{b}}$ would be expected to experience a repulsive interaction from roughly every fourth dislocation that it crosses. Since such a repulsive interaction will be the limiting factor

in the strain relieving process, we choose to study case (i), and hope that the stress fields of arrays A and $\hat{A}$ approximate those of arrays with mixed Burgers vectors. A more sophisticated approach could be adopted, if thought relevant, through use of the techniques developed for alternating arrays in Section 2.6. However, such a refinement is viewed as unwarranted for the preliminary investigation conducted here. Considering, then, the situation in which $\tilde{\mathbf{b}}$ is given by case (i) of Table 3.1 we have calculated, for three values of In fraction x, and as a function of strained layer thickness h, the value of p_a, the largest dislocation spacing for which condition (a) is satisfied. We consider p_a to be the spacing at which the strain relaxation process becomes effectively halted, due to the impediment of propagating dislocations by the dislocations perpendicular to them. The values of x chosen, 0.07, 0.14, and 0.25, correspond to systems studied experimentally by Orders and Usher[54]. These authors used the lattice parameter of the strained layer perpendicular to the interface, which we shall denote $a_\perp$, as their measure of the strain. If we for the moment measure displacement, and hence strain e_{ij}, relative to a configuration in which the layer is subjected to a strain $-e^r_{ij}$ then the perpendicular lattice constant of the layer is given by

$$a_\perp = (1 + e_{22} - e^r_{22})a_l, \tag{3.27}$$

where a_l is the natural lattice constant of the layer, which is assumed to follow Vergaard's law:

$$a_l = 5.65 + 0.40x \text{ Å}.$$

The effect of the fluctuating part of the strain field, which averages to zero, has been neglected. Now, from (2.7)

$$\lambda_{ik}(2e_{2k} - e_{22}\delta_{2k}) = c_{i2kl}e^r_{kl},$$

where $\boldsymbol{\lambda}$ is defined in (2.8). Inversion then yields

$$2e_{2r} - e_{22}\delta_{2r} = (\boldsymbol{\lambda}^{-1})_{ri}c_{i2kl}e^r_{kl}.$$

The system considered here has orthorhombic symmetry, so all elastic constants vanish that have any index appearing an odd number of times. The above equation therfore yields, for the case $r = 2$,

$$e_{22} = \frac{1}{c_{2222}}(c_{1122}e^r_{11} + c_{2222}e^r_{22} + c_{2233}e^r_{33}). \tag{3.28}$$

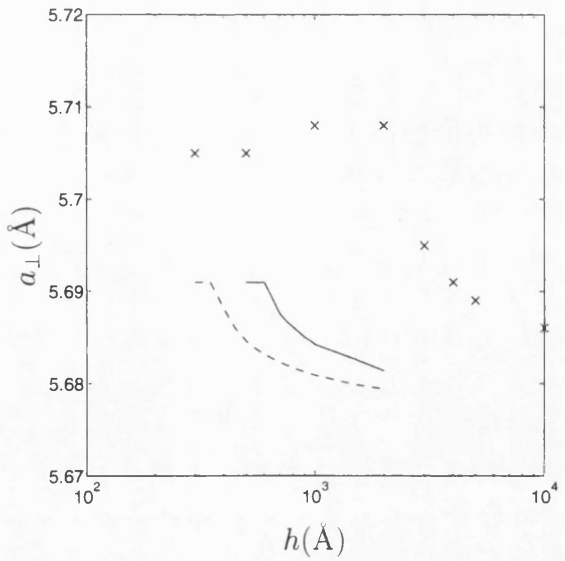

Figure 3.7: Perpendicular lattice constant, $a_\perp$ plotted as a function of layer thickness, h, for an $In_{0.07}Ga_{0.93}As$ layer on GaAs. The solid line gives results obtained using the largest value of p for which condition (a) is satisfied, the broken line gives results obtained using energy minimization, and the crosses mark the experimental results of Orders and Usher.

Now e^r_{ij} is given by (2.47) with e^m_{ij}, $\mathbf{b}$, $\mathbf{n}$, $\hat{\mathbf{b}}$, and $\hat{\mathbf{n}}$ given by (3.9) and (3.11). Therefore,

$$e^r_{11} = f_m + \frac{b_1}{p} = e^r_{33}, \quad e^r_{22} = f_m.$$

In conjunction with (3.28), (3.27) now becomes

$$a_\perp = \left[1 + \frac{c_{1122}}{c_{2222}}\left(f_m + \frac{b_1}{p}\right)\right] a_l. \tag{3.29}$$

Having determined the value of p_a, for given h and f_m, we obtain the resulting values of $a_\perp$ from (3.29). The values of $a_\perp$ are given as a function of h for $x = 0.07$, $x = 0.14$, and $x = 0.25$ in Figs. 3.7, 3.8, and 3.9 respectively. The solid lines give the values obtained using the spacing p_a, the broken lines give the energy minimization values obtained using the spacing p_m, and the crosses mark the experimental results of Orders and Usher. The discrepency between the theoretical and experimental values of $a_\perp$ at small h, when the layers are essentially dislocation-free, may be due to several factors: thermal expansion of the lattice, deviation from Vergaard's law of the natural lattice constant of the alloy, inaccuracies in the interpolation procedure that was used to obtain the elastic constants of the alloys (see Section 3.2), and experimental inaccuracies in determination of the In fraction x. No attempt is made here to

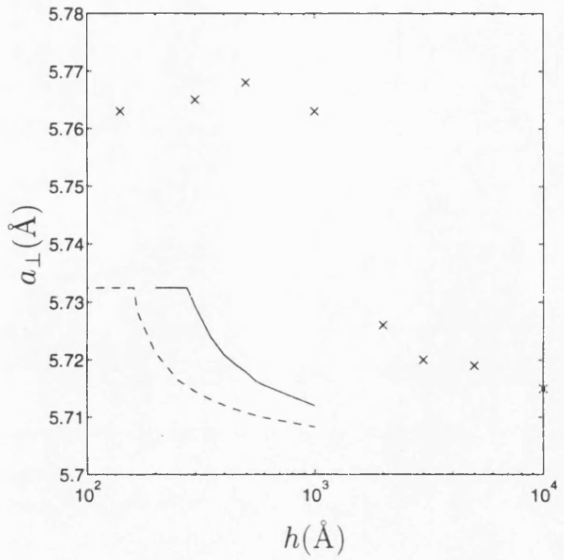

Figure 3.8: As Fig. 3.7 but for an $In_{0.14}Ga_{0.86}As$ layer on GaAs

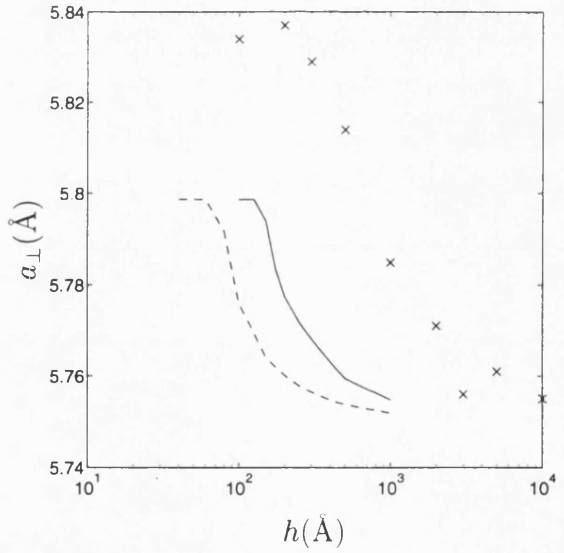

Figure 3.9: As Fig. 3.7 but for an $In_{0.25}Ga_{0.75}As$ layer on GaAs

refine these approximations, because the reconcilement of the discrepency just mentioned will not affect the results that are of interest in the current study. Consider first the critical thickness at which significant mechanical degradation occurs. The experimentally determined values are, from Figs. 3.7 through 3.9, given by, approximately, 2000Å, 900Å, and 200Å for, respectively, $x = 0.07$, 0.14, and 0.25. The values obtained using energy minization are the same as those given in Fig. 2.4 and are 375Å, 160Å, and 80Å. Finally, use of condition (a) in the manner described above yields values of 600Å, 275Å, and 125Å. The sequential approach gives values of effective critical thickness more than one and a half times greater than those obtained using the energy minimization approach. However, even these values are smaller than the experimental ones by a factor of between one and a half and three and a half. Also the theoretical results predict that the extent of mechanical degradation increases faster with layer thickness, as the critical thickness is exceeded, than happens in practice.

The sequential approach to strain relaxation more closely reproduces the experimentally observed behaviour than the energy minimization approach, and yields greater insight into the processes of dislocation propagation. The main reason for this is that the effects of interactions between orthogonal dislocations may be included in the sequential approach. However, the critical thickness at which the onset of the strain relaxation process occurs is severely underestimated even by the sequential approach, and the increase in strain relaxation as the layer thickness is increased above the critical value is overestimated. There are two possible explanations for this. The forces calculated in this chapter used expressions that assumed the existence of perfectly periodic arrays of dislocations. In practice, the dislocations are not periodically arranged, and may be clustered due to the operation of dislocation sources or multiplication mechanisms. If the dislocations in a cluster have like Burgers vector, then the force on the threading arm of a perpendicular dislocation exterted by the cluster will be approximately equal to the force due to a single dislocation, multiplied by the number of dislocations in the cluster. Thus the blocking effect could, in reality, be much larger than suggested by the idealized modelling presented in this chapter. A further reason for the discrepency between theory and experiment is likely to be the lack of any time-dependence in the force model considered here. No allowance has been made for the finite propagation velocity of dislocations across the layer, which may be severely impeded by interaction with perpendicular dislocations, even if ultimately the propagating dislocation is not blocked. Furthermore, in the modelling presented in this chapter, no restriction was placed on the number of dislocations available for the strain relaxation process. However, it is likely that the rate at which dislocations are able to nucleate is a significant factor in determining the extent of strain relaxation. Consideration of this nucleation rate forms the subject matter of

the remainder of this thesis.

Chapter 4

The Stresses due to an Arbitrary Dislocation in a Half-Space

In the preceding chapters it has been illustrated that considerable progress may be made in the study of the propagation and interaction of dislocations using only two-dimensional solutions. However, we shall later want to consider the nucleation of new dislocations in a strained layer. For this purpose, three-dimensional solutions are required. The infinite-body stress fields of many dislocation configurations have been determined explicitly, and the stresses due to an arbitrary dislocation in an infinite body may be written as a line integral around the dislocation, as shown by Mura[52]. By contrast, rather less attention has been paid to three-dimensional dislocation configurations near a free surface. Bacon and Groves[5, 29] have developed a surface-integral representation for the stresses due to an arbitrary dislocation in a half-space. However, this representation is cumbersome and awkward to evaluate. Also, expressions have been obtained for the stresses due to an angular dislocation in a half-space by Comninou and Dundurs[13], from which polygonal dislocation configurations may be constructed.

In this chapter it is demonstrated that the stresses due to an *arbitrary* dislocation in an isotropic half-space may be expressed as a line integral along the dislocation in a manner analogous to the formula of Mura[52] for an infinite medium. The representation obtained is computationally much more convenient than the surface integral representation of Bacon and Groves[5, 29]. For the special case of a dislocation half-line it is demonstrated that the line integral may be evaluated analytically, yielding the stresses in closed form. The dislocation half-line is more convenient than the angular dislocation for the construction of polygonal dislocation configurations, and so the study of such configurations may be carried through with

more simplicity than was previously possible.

4.1 Basic Preliminaries

The problem will be to evaluate the stresses σ_{ij} due to a dislocation of Burgers vector $\mathbf{b}$ in a uniform, isotropic, elastic half-space. To make the displacement field $\mathbf{u}$ due to the dislocation single-valued, a surface S is defined across which the tractions are continuous, while the displacement has a discontinuity $\mathbf{b}$. It is well known that the stresses are continuous across S if $\mathbf{b}$ is constant, and so only depend upon the boundary of S, namely the dislocation line. The stresses are related to the displacement in the usual way, via the tensor of elastic moduli:

$$\sigma_{ij} = c_{ijkl} u_{k,l}.$$

The stresses must obey the equilibrium equations, which, cast in terms of displacement, become

$$c_{ijkl} u_{k,lj} = 0, \tag{4.1}$$

since body forces are assumed to be absent. By casting the problem in terms of displacement, compatability relations between the strains are automatically satisfied. The problem is thus to determine the displacement $\mathbf{u}$ in the body B satisfying (4.1) together with the boundary conditions

$$\mathbf{u} \to 0 \text{ as } |\mathbf{x}| \to \infty \text{ within } B,$$

$$[u_i] = b_i \text{ and } [c_{ijkl} n_j u_{k,l}] = 0 \text{ across } S, \tag{4.2}$$

$$c_{ijkl} n_j u_{k,l} = 0 \text{ on } \partial B, \tag{4.3}$$

where $[f]$ represents the jump in f across S, and n_j are the components of the normal to S or ∂B. Condition (4.3) expresses the condition of zero traction on the outer surface ∂B of the body B within which the dislocation resides.

Define the normal $\mathbf{m}$ to the surface S so that $\mathbf{b}$ is the jump in displacement as S is crossed in the direction of $\mathbf{m}$. (This is the Burgers vector convention described and used in Section 3.1.) Then the displacement field specified by (4.1) through (4.3) may be expressed as[71]

$$u_m(\mathbf{x}') = \int_S b_i c_{ijkl} G_{mk,l}(\mathbf{x}, \mathbf{x}') m_j dS, \tag{4.4}$$

where $G_{mk}(\mathbf{x},\mathbf{x}')$ is the $mk-$component of the Green's tensor for the body B, obeying

$$c_{ijkl}G_{mk,lj}(\mathbf{x},\mathbf{x}') + \delta_{im}\delta(\mathbf{x}-\mathbf{x}') = 0 \tag{4.5}$$

together with homogeneous boundary conditions on the boundary of B:[1]

$$c_{ijkl}n_j G_{mk,l}(\mathbf{x},\mathbf{x}') = 0 \text{ for } \mathbf{x} \in \partial B.$$

In all of the above, differentiation is with respect to elements of the first vector variable. The stresses are easily obtained from (4.4) by differentiation:

$$\sigma_{pq}(\mathbf{x}') = \int_S b_s c_{srkl} c_{pqmj} \frac{\partial}{\partial x'_j} G_{mk,l}(\mathbf{x},\mathbf{x}') m_r dS. \tag{4.6}$$

The above representation is valid for a dislocation in any elastic body, provided that the Green's tensor appropriate for that body is used. In particular, (4.6) with the half-space Green's tensor obtained by Mindlin[51] is in essence the representation used by Bacon and Groves[5, 29] and, more recently, by Beltz and Freund[7]. In the general case, (4.6) is a cumbersome representation, especially for use in energy evaluations, where a further integral over the surface S must be evaluated.

For an infinite medium, translational symmetry ensures that G_{mk} is a function of $\mathbf{x}-\mathbf{x}'$, and in this case Stokes' theorem may be employed to transform the surface integral of (4.6) into a line integral around the dislocation line C, which bounds S:

$$\sigma_{pq}(\mathbf{x}) = \oint_C b_s S^{\infty}_{pqrs}(\mathbf{x}-\mathbf{x}')dx'_r, \tag{4.7}$$

where

$$S^{\infty}_{pqrs}(\mathbf{x}-\mathbf{x}') = c_{snkl}c_{pqmj}\epsilon_{njr}G^{\infty}_{mk,l}(\mathbf{x}-\mathbf{x}'). \tag{4.8}$$

Here G^{∞}_{mk} represents the infinite body Green's tensor, and ϵ_{ijk} is the completely antisymmetric unit tensor of order three. Equation (4.7) was first derived by Mura[52]. It should be emphasized that Mura's representation cannot be made valid for a half-space simply by replacing the infinite-body Green's tensor by the half-space Green's tensor, since the latter does not possess all of the required symmetries. However, a line-integral representation may be constructed by a different method, as is now demonstrated.

[1] If B is *finite*, a modification is needed to ensure overall equilibrium. The simplest would be to require zero displacement on some part of ∂B. However, this is not an issue for a half-space.

The kernel $S^{\infty}_{pqrs}(\mathbf{x}-\mathbf{x}')$ of Mura's formula may be viewed as being the $pq-$stress at $\mathbf{x}$ produced by a line element of dislocation lying in the r-direction at $\mathbf{x}'$, in an infinite body, with unit Burgers vector in the s-direction. The tensor S^{∞}_{pqrs} thus represents nine such stress sources. If we now consider a half-space, each of these sources produces tractions on the free surface, which must be annulled in order to satisfy the free-surface boundary condition. The programme is thus to find the nine 'image' stress fields, denoted $S^{I}_{pqrs}(\mathbf{x},\mathbf{x}')$, such that the stress field due to a dislocation loop in a half-space may be written

$$\sigma_{pq}(\mathbf{x}) = \oint_C b_s \left(S^{\infty}_{pqrs}(\mathbf{x}-\mathbf{x}') + S^{I}_{pqrs}(\mathbf{x},\mathbf{x}')\right) dx'_r. \tag{4.9}$$

The jump condition on the displacement, (4.2), is automatically satisfied by the term involving S^{∞}_{pqrs}, as for an infinite body. The term S^{I}_{pqrs} must therefore introduce no new singularities in the half-space: the rôle of this term is merely to annul the free surface tractions.

4.2 Construction of the Line-Integral Representation

The mathematical problem to be addressed is now formally defined. Let the half-space occupy $x_3 > 0$. We seek nine continuous displacement fields $\mathbf{u}_{rs}$ that are equilibrium fields, so that

$$c_{ijkl}u_{krs,lj}(\mathbf{x},\mathbf{x}') = 0. \tag{4.10}$$

The displacement fields $\mathbf{u}_{rs}$ must also satisfy the boundary conditions

$$\mathbf{u}_{rs}(\mathbf{x},\mathbf{x}') \to 0 \; as \; |\mathbf{x}| \to \infty, \quad \text{within the body,}$$

$$c_{p3kl}u_{krs,l}(\mathbf{x},\mathbf{x}')\big|_{x_3=0} = -S^{\infty}_{p3rs}(\mathbf{x}-\mathbf{x}')\big|_{x_3=0}.$$

Then the source-image tensor that makes representation (4.9) valid is

$$S^{I}_{pqrs}(\mathbf{x},\mathbf{x}') = c_{pqkl}u_{krs,l}(\mathbf{x},\mathbf{x}').$$

The problem is most conveniently solved by Fourier transforming with respect to x_1 and x_2. Exploiting the translational symmetry of the system, we consider a source lying on the line $x'_1 = x'_2 = 0$. Then defining

$$\hat{\mathbf{u}}(\boldsymbol{\xi}, x_3, x'_3) = \int\int dx_1 dx_2 \mathbf{u}(\mathbf{x},(0,0,x'_3))e^{\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{y}},$$

where $\boldsymbol{\xi} = (\xi_1, \xi_2)$ and $\mathbf{y} = (x_1, x_2)$, the equilibrium equations (4.10) transform to

$$c_{i3k3}\frac{\partial^2 \hat{u}_{krs}}{\partial x_3^2} - \mathrm{i}\xi_\alpha(c_{i\alpha k3} + c_{i3k\alpha})\frac{\partial \hat{u}_{krs}}{\partial x_3} - \xi_\alpha \xi_\beta c_{i\alpha k\beta}\hat{u}_{krs} = 0. \tag{4.11}$$

Here the convention has been adopted that Greek subscripts range over 1, 2, whereas Roman subscripts range over 1, 2, 3. The body is assumed to be elastically isotropic so that

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}),$$

where λ and μ are the Lamé constants for the body. In this case the general solution of (4.11) that decays as $x_3 \to \infty$ may be written

$$\hat{\mathbf{u}}_{rs}(\boldsymbol{\xi}, x_3) = A_{1rs}\begin{pmatrix} \eta_1 \\ \eta_2 \\ -\mathrm{i} \end{pmatrix} \mathrm{e}^{-|\xi|x_3} + A_{2rs}\begin{pmatrix} \eta_2 \\ -\eta_1 \\ 0 \end{pmatrix} \mathrm{e}^{-|\xi|x_3}$$

$$+ A_{3rs}\left[\begin{pmatrix} \eta_1 \\ \eta_2 \\ -\mathrm{i} \end{pmatrix} x_3 + \frac{1}{|\xi|}\left(\frac{\lambda + 3\mu}{\lambda + \mu}\right)\begin{pmatrix} 0 \\ 0 \\ -\mathrm{i} \end{pmatrix}\right] \mathrm{e}^{-|\xi|x_3}, \tag{4.12}$$

where the normalized reciprocal vector has been introduced:

$$\eta = \xi / |\xi| .$$

The constants A_{qrs} are determined by the zero-traction condition at $x_3 = 0$ for each source. In the isotropic case considered here, the problem is simplified by noting that S^I_{ijkl} is isotropic with respect to rotations in the x_1-x_2 plane, so that

$$R_{pi}R_{qj}R_{rk}R_{sl}S^I_{ijkl}(\mathbf{x}, x_3') = S^I_{pqrs}(\mathbf{Rx}, x_3'),$$

where $\mathbf{R}$ is the matrix representing any rotation about the x_3-axis. On transforming it is found that

$$R_{pi}R_{qj}R_{rk}R_{sl}\hat{S}^I_{ijkl}(\eta, |\xi|, x_3, x_3') = \hat{S}^I_{pqrs}(\tilde{\mathbf{R}}\eta, |\xi|, x_3, x_3'), \tag{4.13}$$

where $\tilde{\mathbf{R}}$ is the 2×2 matrix given by $\tilde{R}_{\alpha\beta} = R_{\alpha\beta}$. In particular, for each η we may choose

$$\mathbf{R} = \begin{pmatrix} \eta_1 & \eta_2 & 0 \\ -\eta_2 & \eta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.14}$$

Then

$$\tilde{\mathbf{R}}\boldsymbol{\eta} = (1,0) = \mathbf{e},$$

say, and (4.13) yields

$$\hat{S}^I_{ijkl}(\boldsymbol{\eta}, |\boldsymbol{\xi}|, x_3, x_3') = R_{pi}R_{qj}R_{rk}R_{sl}\hat{S}_{pqrs}(\mathbf{e}, |\boldsymbol{\xi}|, x_3, x_3'). \tag{4.15}$$

Thus only one reciprocal vector direction, for convenience the direction $\boldsymbol{\eta} = (1,0)$, need be considered, with the general case following trivially from (4.15) with $\mathbf{R}$ given by (4.14). This observation yields a considerable simplification of the algebra. When $\boldsymbol{\eta} = (1,0)$ the stresses engendered by the displacements of (4.12) are

$$\begin{aligned}
\hat{S}^I_{11rs} &= -2\mu\mathrm{i}\left[|\boldsymbol{\xi}| A_{1rs} + \left(|\boldsymbol{\xi}| x_3 - \frac{\lambda}{\lambda+\mu}\right) A_{3rs}\right] \mathrm{e}^{-|\boldsymbol{\xi}|x_3}, \\
\hat{S}^I_{12rs} &= \mu\mathrm{i}\,|\boldsymbol{\xi}| A_{2rs}\mathrm{e}^{-|\boldsymbol{\xi}|x_3}, \\
\hat{S}^I_{13rs} &= -2\mu\left[|\boldsymbol{\xi}| A_{1rs} + \left(|\boldsymbol{\xi}| x_3 + \frac{\mu}{\lambda+\mu}\right) A_{3rs}\right] \mathrm{e}^{-|\boldsymbol{\xi}|x_3}, \\
\hat{S}^I_{22rs} &= \frac{2\mu\lambda\mathrm{i}}{\lambda+\mu} A_{3rs}\mathrm{e}^{-|\boldsymbol{\xi}|x_3}, \\
\hat{S}^I_{23rs} &= \mu\,|\boldsymbol{\xi}| A_{2rs}\mathrm{e}^{-|\boldsymbol{\xi}|x_3}, \\
\hat{S}^I_{33rs} &= 2\mu\mathrm{i}\left[|\boldsymbol{\xi}| A_{1rs} + \left(|\boldsymbol{\xi}| x_3 + \frac{\lambda+2\mu}{\lambda+\mu}\right) A_{3rs}\right] \mathrm{e}^{-|\boldsymbol{\xi}|x_3}.
\end{aligned} \tag{4.16}$$

The surface tractions at $x_3 = 0$ may be written

$$\hat{S}^I_{i3rs}\Big|_{x_3=0} = M_{ij}A_{jrs},$$

where

$$\mathbf{M} = \begin{pmatrix} -2\mu\,|\boldsymbol{\xi}| & 0 & -2\mu^2/(\lambda+\mu) \\ 0 & \mu\,|\boldsymbol{\xi}| & 0 \\ 2\mu\mathrm{i}\,|\boldsymbol{\xi}| & 0 & 2\mu\mathrm{i}(\lambda+2\mu)/(\lambda+\mu) \end{pmatrix}.$$

The tractions given above must annul the source tractions on $x_3 = 0$. Therefore,

$$A_{qrs} = -(\mathbf{M}^{-1})_{qp}\hat{S}^{\infty}_{p3rs}\Big|_{x_3=0}. \tag{4.17}$$

The transformed image stresses are thus defined completely, once the transformed source stresses are known. From (4.8),

$$\hat{S}^{\infty}_{p3rs} = -i\xi_\alpha c_{snk\alpha}c_{p3mj}\epsilon_{njr}\hat{G}^{\infty}_{mk} + c_{snk3}c_{p3mj}\epsilon_{njr}\frac{\partial}{\partial x_3}\hat{G}^{\infty}_{mk}. \tag{4.18}$$

Hence the two-dimensional transforms of the components of the infinite-body Green's tensor are required, together with their derivatives with respect to x_3. The tractions (4.18) due to a source at $x_3' > 0$ are to be evaluated at $x_3 = 0$. Therefore, the transformed Green's tensor, $\hat{G}^{\infty}_{mk}(\boldsymbol{\xi}, x_3 - x_3')$ is required when $x_3 - x_3' < 0$. In this case the transformed Green's tensor has components (see Appendix B)

$$\begin{aligned}
\hat{G}^{\infty}_{\alpha\beta} &= \delta_{\alpha\beta}\frac{\mathrm{e}^{-|\xi|(x_3'-x_3)}}{2\mu\,|\xi|} - \frac{1}{\mu}\left(\frac{\lambda+\mu}{\lambda+2\mu}\right)\eta_\alpha\eta_\beta(1+|\xi|\,(x_3'-x_3))\frac{\mathrm{e}^{-|\xi|(x_3'-x_3)}}{4\,|\xi|},\\
\hat{G}^{\infty}_{\alpha 3} &= -\frac{1}{\mu}\left(\frac{\lambda+\mu}{\lambda+2\mu}\right)\eta_\alpha \mathrm{i}(x_3'-x_3)\frac{\mathrm{e}^{-|\xi|(x_3'-x_3)}}{4},\\
\hat{G}^{\infty}_{33} &= \frac{\mathrm{e}^{-|\xi|(x_3'-x_3)}}{2\mu\,|\xi|} - \frac{1}{\mu}\left(\frac{\lambda+\mu}{\lambda+2\mu}\right)(1-|\xi|\,(x_3'-x_3))\frac{\mathrm{e}^{-|\xi|(x_3'-x_3)}}{4\,|\xi|}.
\end{aligned} \tag{4.19}$$

The transformed image stresses are now completely defined through (4.15)-(4.19). Lengthy but straightforward algebra (performed using the symbolic manipulation programme REDUCE) reveals that when $\eta = (1,0)$, the transformed image stresses take the form

$$\hat{S}^{I}_{ijkl} = K_{ijklm}\,|\xi|^{m-1}\,\mathrm{e}^{-|\xi|(x_3+x_3')}.$$

Summation on repeated *lower case* Roman indices takes place over the values 1, 2, 3 of that index. The coefficients K_{ijklm} depend on x_3, x_3' and the elastic constants. From (4.15), the general transformed stresses are given by

$$\hat{S}^{I}_{ijkl}(\boldsymbol{\xi}, x_3, x_3') = R_{pi}R_{qj}R_{rk}R_{sl}K_{pqrsm}\,|\xi|^{m-1}\,\mathrm{e}^{-|\xi|(x_3+x_3')}. \tag{4.20}$$

It is clear from (4.14) that

$$R_{pi}R_{qj}R_{rk}R_{sl} = \eta_1^N\eta_2^P,$$

where N, P are some non-negative integers satisfying $0 \leq N + P \leq 4$. We may therefore introduce coefficients $W_{ijklmNP}$, dependent upon x_3, x_3', and the elastic constants, such that

$$R_{pi}R_{qj}R_{rk}R_{sl}K_{pqrsm} = W_{ijklmNP}\eta_1^N\eta_2^P. \tag{4.21}$$

The convention has been adopted that repeated *upper case* Roman indices sum over the values 0, 1, 2, 3, 4 of that index. With the notation of (4.21), (4.20) becomes

$$\hat{S}^I_{ijkl}(\boldsymbol{\xi}, x_3, x_3') = W_{ijklmNP}\eta_1^N\eta_2^P\,|\boldsymbol{\xi}|^{m-1}\,\mathrm{e}^{-|\xi|(x_3+x_3')}.$$

Inversion, using polar coordinates, yields

$$\begin{aligned} S^I_{ijkl}(x_1 - x_1', x_2 - x_2', x_3, x_3') &= \frac{1}{4\pi^2}W_{ijklmNP} \\ &\quad \times \oint_{|\eta|=1} ds\eta_1^N\eta_2^P \int_0^\infty d\,|\boldsymbol{\xi}|\,|\boldsymbol{\xi}|^m\,\mathrm{e}^{-|\xi|[x_3+x_3'+\mathrm{i}\eta.(\mathbf{y}-\mathbf{y}')]}, \end{aligned} \tag{4.22}$$

where $\mathbf{y}=(x_1, x_2)$, $\mathbf{y}' = (x_1', x_2')$. The radial integral may be performed at sight to yield

$$S^I_{ijkl} = \frac{1}{4\pi^2}W_{ijklmNP}\left(-\frac{\partial}{\partial x_3}\right)^m \oint_{|\eta|=1} ds \frac{\eta_1^N\eta_2^P}{x_3 + x_3' + \mathrm{i}\eta.(\mathbf{y}-\mathbf{y}')}.$$

Recalling that $0 \leq N + P \leq 4$ ($W_{ijklmNP} = 0$ for $N + P > 4$), it is clear that the terms $\eta_1^N\eta_2^P$ may be written as linear combinations of terms $(\eta_1 \pm \mathrm{i}\eta_2)^Q$, $0 \leq Q \leq 4$. Therefore, we may introduce coefficients $\omega^{\pm}_{NPQ}$ such that

$$\eta_1^N\eta_2^P = \omega^+_{NPQ}(\eta_1 + \mathrm{i}\eta_2)^Q + \omega^-_{NPQ}(\eta_1 - \mathrm{i}\eta_2)^Q.$$

Then

$$S^I_{ijkl} = \frac{1}{4\pi^2}W_{ijklmNP}\left(-\frac{\partial}{\partial x_3}\right)^m \oint_{|\eta|=1} ds \frac{\omega^+_{NPQ}(\eta_1 + \mathrm{i}\eta_2)^Q + \omega^-_{NPQ}(\eta_1 - \mathrm{i}\eta_2)^Q}{x_3 + x_3' + \mathrm{i}\eta.(\mathbf{y}-\mathbf{y}')}.$$

Defining

$$J^{\pm}_Q = \oint_{|\eta|=1} ds \frac{(\eta_1 \pm \mathrm{i}\eta_2)^Q}{x_3 + x_3' + \mathrm{i}\eta.(\mathbf{y}-\mathbf{y}')}, \tag{4.23}$$

the image stresses may be written in the compact form

$$S^I_{ijkl} = \frac{1}{4\pi^2}W_{ijklmNP}\left(-\frac{\partial}{\partial x_3}\right)^m (\omega^+_{NPQ}J^+_Q + \omega^-_{NPQ}J^-_Q).$$

The integrals of (4.23) are easily evaluated using Cauchy's theorem. This is carried out in detail in Appendix C, yielding the results

$$J_Q^{\pm} = \frac{2\pi}{R}\left(\frac{[\mathrm{i}(x_1 - x_1') \mp (x_2 - x_2')](x_3 + x_3' - R)}{r^2}\right)^Q, \tag{4.24}$$

where

$$R^2 = (x_1 - x_1')^2 + (x_2 - x_2')^2 + (x_3 + x_3')^2,$$

and

$$r^2 = (x_1 - x_1')^2 + (x_2 - x_2')^2.$$

The term required to complete the integral representation, (4.9), for the stresses due to a dislocation in a half-space has thus been obtained in closed form.

4.3 The Stresses due to a Dislocation Half-Line in a Half-Space

The stresses due to a dislocation half-line in a whole-space have been documented by Li[48] and so only the 'image' terms arising from the presence of a free surface are discussed here.

Consider a half-line L parameterized by

$$\mathbf{x}(t) = \mathbf{a} + \boldsymbol{\alpha} t; \quad 0 \leq t < \infty.$$

Then the contribution to the stresses from the free-surface correction, which must be added to the infinite body stresses of Li[48], is

$$\sigma_{ij}^I(\mathbf{x}; \mathbf{a}, \boldsymbol{\alpha}) = \int_0^\infty dt b_l \alpha_k S_{ijkl}^I(\mathbf{x}, \mathbf{a} + \boldsymbol{\alpha} t). \tag{4.25}$$

The integral is performed using the Fourier integral representation for S_{ijkl}^I, (4.22), in (4.25). It is first necessary to make the dependence on x_3' in (4.22) completely explicit. It may be demonstrated that the coefficients $W_{ijklmNP}$ are of the form

$$W_{ijklmNP} = W_{ijklmNP}^0 + x_3 W_{ijklmNP}^1 + x_3' W_{ijklmNP}^2 + x_3' x_3 W_{ijklmNP}^3,$$

where the constants $W_{ijklmNP}^r$, $r = 0, 1, 2, 3$ depend only upon the elastic constants. For our

purposes it is only necessary to make the dependence upon x_3' explicit, and so we shall write

$$W_{ijklmNP} = U_{ijklmNP} + x_3' V_{ijklmNP}, \tag{4.26}$$

where the constants $U_{ijklmNP}$ and $V_{ijklmNP}$ may depend on x_3 and the elastic constants, but not upon x_3'. Inserting (4.22) in (4.25), and making use of (4.26) yields

$$\sigma_{ij}^I = \frac{1}{4\pi^2} b_l \alpha_k \int_0^\infty dt\, [U_{ijklmNP} + (a_3 + \alpha_3 t) V_{ijklmNP}]$$

$$\times \oint_{|\eta|=1} ds \eta_1^N \eta_2^P \int_0^\infty d\,|\xi|\,|\xi|^m \, \mathrm{e}^{-|\xi|[x_3+a_3+\mathrm{i}\eta_1(x_1-a_1)+\mathrm{i}\eta_2(x_2-a_2)+t(\alpha_3-\mathrm{i}\eta_1\alpha_1-\mathrm{i}\eta_2\alpha_2)]}.$$

Performing the integral with respect to t first, followed by the integral with respect to $|\xi|$ yields

$$\begin{aligned}
\sigma_{ij}^I &= \frac{1}{4\pi^2} b_l \alpha_k (U_{ijklmNP} + a_3 V_{ijklmNP}) \left(-\frac{\partial}{\partial x_3}\right)^{m-1} \\
&\quad \times \oint_{|\eta|=1} ds \frac{\omega^+_{NPQ}(\eta_1 + \mathrm{i}\eta_2)^Q + \omega^-_{NPQ}(\eta_1 - \mathrm{i}\eta_2)^Q}{(\alpha_3 - \mathrm{i}\eta_1\alpha_1 - \mathrm{i}\eta_2\alpha_2)[x_3 + a_3 + \mathrm{i}\eta_1(x_1 - a_1) + \mathrm{i}\eta_2(x_2 - a_2)]} \\
&\quad + \frac{1}{4\pi^2} b_l \alpha_k \alpha_3 V_{ijklmNP} \left(-\frac{\partial}{\partial \alpha_3}\right) \left(-\frac{\partial}{\partial x_3}\right)^{m-2} \\
&\quad \times \oint_{|\eta|=1} ds \frac{\omega^+_{NPQ}(\eta_1 + \mathrm{i}\eta_2)^Q + \omega^-_{NPQ}(\eta_1 - \mathrm{i}\eta_2)^Q}{(\alpha_3 - \mathrm{i}\eta_1\alpha_1 - \mathrm{i}\eta_2\alpha_2)[x_3 + a_3 + \mathrm{i}\eta_1(x_1 - a_1) + \mathrm{i}\eta_2(x_2 - a_2)]}.
\end{aligned} \tag{4.27}$$

For the integral with respect to t to be convergent it is necessary that $\alpha_3 > 0$, that is, the dislocation half-line must run away from the free surface into the half-space. However, the stresses due to a half-line parallel to the free surface may be deduced from (4.27) by taking the limiting case of $\alpha_3 = 0$, which yields a finite result. By introducing the notation

$$H_Q^\pm = \oint_{|\eta|=1} ds \frac{(\eta_1 \pm \mathrm{i}\eta_2)^Q}{(\alpha_3 - \mathrm{i}\eta_1\alpha_1 - \mathrm{i}\eta_2\alpha_2)[x_3 + a_3 + \mathrm{i}\eta_1(x_1 - a_1) + \mathrm{i}\eta_2(x_2 - a_2)]},$$

(4.27) may be written succinctly as

$$\begin{aligned}
\sigma_{ij}^I &= \frac{1}{4\pi^2} b_l \alpha_k \left[(U_{ijklmNP} + a_3 V_{ijklmNP}) \left(-\frac{\partial}{\partial x_3}\right)^{m-1} + \alpha_3 V_{ijklmNP} \left(-\frac{\partial}{\partial \alpha_3}\right) \left(-\frac{\partial}{\partial x_3}\right)^{m-2} \right] \\
&\quad \times (\omega^+_{NPQ} H_Q^+ + \omega^-_{NPQ} H_Q^-).
\end{aligned}$$

It may be demonstrated that when $m = 1$, $V_{ijklmNP} = 0$, so no special interpretation of $(-\partial/\partial x_3)^{m-2}$ is required. The integrals $H_Q^\pm$ are easily evaluated using Cauchy's theorem, as

illustrated in Appendix D. Introducing the notation

$$f_\pm = \frac{\alpha_3 \pm \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}}{\mathrm{i}\alpha_1 + \alpha_2},$$

$$g_\pm = \frac{-(x_3 + a_3) \pm \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + (x_3 + a_3)^2}}{\mathrm{i}(x_1 - a_1) + (x_2 - a_2)},$$

the results may be written

$$H_Q^+ = \frac{-8\pi}{(\mathrm{i}\alpha_1 + \alpha_2)[\mathrm{i}(x_1 - a_1) + x_2 - a_2]} \left[\frac{(f_-)^{Q+1}}{(f_- - f_+)(f_- - g_+)(f_- - g_-)} \right.$$

$$\left. + \frac{(g_+)^{Q+1}}{(g_+ - f_+)(g_+ - f_-)(g_+ - g_-)} \right], \tag{4.28}$$

and

$$H_Q^- = H_Q^+\Big|_{\alpha_2 = -\alpha_2,\ x_2 - a_2 = a_2 - x_2} = (-1)^Q \overline{H_Q^+} \quad \text{(no summation)}. \tag{4.29}$$

The stresses due to a dislocation half-line in a half-space are thus obtained completely explicitly. The stresses due to a dislocation segment follow trivially: if we wish to know the stress field of the dislocation segment joining $\mathbf{a}$ to $\mathbf{c}$, where $c_3 \geq a_3$, then defining $\boldsymbol{\alpha} = \mathbf{c} - \mathbf{a}$ we may simply write the stresses of the segment as

$$\sigma_{ij}(\mathbf{x}) = \sigma_{ij}(\mathbf{x}; \mathbf{a}, \boldsymbol{\alpha}) - \sigma_{ij}(\mathbf{x}; \mathbf{c}, \boldsymbol{\alpha}).$$

4.4 Special Considerations for a Surface Half-Loop

A case of practical interest is that of nucleation of a half-loop at the free surface of a strained layer. In this case, the line-integral representation, as constructed in Section 4.2, requires the integral, along a segment of dislocation lying in the free surface, of a difference of terms each of which is singular, but whose difference is finite. In order to circumvent this numerical inconvenience, a modified construction of the line-integral representation is desirable. By contrast with (4.9) we now seek a term $S^{I\star}_{pqrs}$ such that the representation

$$\sigma_{pq} = \oint_C b_s \left(S^{\infty\star}_{pqrs}(\mathbf{x}, \mathbf{x}') + S^{I\star}_{pqrs}(\mathbf{x}, \mathbf{x}') \right) dx'_r \tag{4.30}$$

gives the stresses due to a half-loop C in a half-space, where

$$S^{\infty\star}_{pqrs}(\mathbf{x}, \mathbf{x}') = S^{\infty}_{pqrs}(\mathbf{x} - \mathbf{x}') - S^{\infty}_{pqrs}(\mathbf{x} + \mathbf{x}'). \tag{4.31}$$

In (4.30) the origin of coordinates has been chosen to coincide with the midpoint of the line joining the points of intersection of the loop with the free surface. The singular terms are now all incorporated within the source term, $S^{\infty\star}_{pqrs}$. Along the interfacial dislocation segment this source term is not zero (in fact it is singular there); however, the *integral* of the source term along the interfacial segment vanishes. Therefore, the integral along the interfacial segment of the image term, $S^{I\star}_{pqrs}$, also vanishes. Thus the interfacial segment of dislocation does not contribute to the integral in (4.30), which may be performed just around the half-loop. By this method, awkward numerical singularities are avoided.

Since the problem is linear, the contribution to the image stresses from the two terms on the righthand side of (4.31) may be evaluated separately and then summed. The contribution from $S^{\infty}_{pqrs}(\mathbf{x}-\mathbf{x}')$ was considered explicitly in Sections 4.2 and 4.3, so only the term $S^{\infty}_{pqrs}(\mathbf{x}+\mathbf{x}')$ requires further attention. It is found that the transformed Green's tensor $\hat{G}^{\infty}_{mk}(\boldsymbol{\xi},z)$ is now required when $z = x_3 + x_3' > 0$. It is shown in Appendix B that the components in this case are obtained from the components of (4.19) by making the replacement $x_3' - x_3 \mapsto x_3' + x_3$ and taking the complex conjugate. Derivation of the contribution to the *transformed* image stresses now proceeds precisely as in Section 4.2. Only the coefficients $W^{r}_{ijklmNP}$, r=0, 1, 2, 3, are altered. On inversion it should be noted that x_1, x_1', x_2, x_2' now appear in the combinations $x_1 + x_1'$, $x_2 + x_2'$, so the new functions $H^{\pm}_{Q}$, $J^{\pm}_{Q}$ are obtained by making the replacements $x_\beta' \mapsto -x_\beta'$, $a_\beta \mapsto -a_\beta$, $\alpha_\beta \mapsto -\alpha_\beta$, $\beta = 1,\ 2$, in (4.24), (4.28), and (4.29).

It was mentioned that in (4.30) the origin of coordinates should be chosen to be the midpoint of the line joining the two points of intersection of the half-loop with the free surface. If this condition is satisfied then

$$\oint_C b_s S^{\infty\star}_{pqrs}(\mathbf{x},\mathbf{x}')dx_r' = \oint_{C_0} b_s S^{\infty}_{pqrs}(\mathbf{x}-\mathbf{x}')dx_r',$$

where C_0 is the whole-loop obtained by completing the half-loop C in the following manner:

$$\{\mathbf{x} \in C_0\} = \{\mathbf{x} \mid \mathbf{x} \in C \text{ or } -\mathbf{x} \in C\}.$$

Thus the first term in the integral representation, (4.30), gives the stresses due to the completed loop in an infinte body, and the second term provides the half-space correction.

The formulae obtained here will be exploited in the next chapter.

Chapter 5

Aspects of Nucleation

5.1 General Considerations

In this chapter we make predictions of the rate at which dislocations will nucleate in a strained-layer system. By considering nucleation rates, a time-dependent element is introduced into our study of the strain relaxation process, which was absent in Chapter 3. Determination of this time dependence is of great practical importance, because it determines the timescale over which dislocations are introduced during both processing and use, and therefore determines the effective lifetime of a device. In general, the nucleation process is extremely complicated, with many different possible sites and mechanisms. No more than a partial description should be expected here. However, we shall aim to explain qualitatively, and with partial quantification, the variation in the type of dislocation distribution observed as the lattice mismatch is increased in both the InGaAs/GaAs(100) and GeSi/Si(100) strained-layer systems. The variation is the same in the two systems, and is now described. The most direct experimental support for the description that follows may be found in the work of Kvam *et al.* and Chang *et al.*[44, 12], who studied structures with a range of mismatch strains in the same set of experiments. Further supporting evidence is found in a number of other experimental papers[37, 44, 20, 32, 33, 67, 36, 60, 43].

At low mismatch strains ($f_m < 0.015$) the threading segment density is rather low, and relaxation occurs via the introduction of long 60° misfit dislocations, irregularly spaced in the interface and often collected in clusters. As the lattice mismatch increases above 0.015 to 0.025 the distribution changes markedly: the threading segment density increases rapidly by about two orders of magnitude, the dislocations are observed to be very much shorter than at low

mismatch strains and to be randomly spaced not clustered, and, rather than 60° dislocations, 90° dislocations are observed to be the norm. (90° dislocations run in the same ⟨110⟩ interface directions as 60° dislocations, but have Burgers vector along the *interfacial* ⟨110⟩/2 lattice vector that is perpendicular to the dislocation line: the terminology 60° and 90° refers to the angle made between the dislocation line and its Burgers vector.) To our knowledge the mismatch range 0.035 to 0.07 has not been studied systematically. However, studies do exist of very highly mismatched systems for which f_m ranges between 0.07 and 0.14. In these systems almost perfectly periodic arrays of 90° dislocations have been observed. In this chapter we shall concentrate on modelling and comparing with experiment the variation in dislocation distribution for $f_m < 0.35$. In the very highly strained systems that have been studied, mechanical degradation occurs after just a few mononlayers of growth and has been associated with dislocation formation around the edges of three-dimensional growth features (i.e.'islands'). It is not clear how applicable the continuum model of a dislocation is to such a situation, or how appropriate it is to model the free surface as flat. For these reasons, in this chapter we shall restrict our attention to the mismatch range $f_m < 0.035$. In these systems, we are confident that the type of modelling pursued here is appropriate.

The formation of dislocations via loop nucleation is illustrated in Fig. 5.1 where a pair of propagating threading segments is generated (a) from the nucleation of a loop within the layer, and (b) from the nucleation of a half loop at the free surface. In most cases of interest the energy of the system will increase with loop size when the loop size is small, because the self-energy of the loop will outweigh the energy reduction due to strain relief. Above a critical size the total energy will decrease as the loop expands, as the energy reduction due to strain relief begins to dominate. Let W_{act} denote the maximum loop energy . (That is, the maximum energy measured relative to the energy of the system before introduction of the loop.) This energy, which we shall call the loop activation energy, must be supplied thermally if a loop is to nucleate, thereby generating a new dislocation. Kamat and Hirth[42] have provided a simple expression relating the nucleation rate, J, to the loop activation energy:

$$J \approx 10^{36} \exp(-W_{act}/kT) \mathrm{cc}^{-1}\mathrm{s}^{-1}, \tag{5.1}$$

where k is Boltzmann's constant and T is the absolute temperature. Equation (5.1) is derived from a treatment essentially transported from the study of droplet formation in supersaturated vapours[18]. Inclusion of the Arrhenius type exponential is quite standard, but great uncertainty exists concerning the choice of pre-exponential factor. Kamat and Hirth estimate that the choice

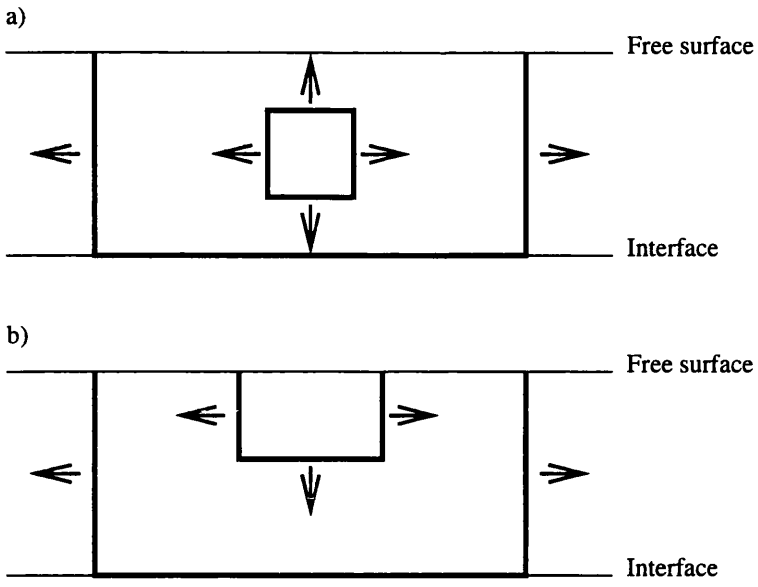

Figure 5.1: Schematic illustration of the side-on view of dislocation formation via loop nucleation in a strained layer: a) nucleation of a loop buried in the layer and b) nucleation of a surface half-loop. In both cases the glide plane will, in fact, be inclined to the free surface.

of 10^{36} used here is likely to be rather an over-estimate. The significance of the pre-exponential is discussed later. In this chapter we use (5.1) to predict the rates of nucleation at various sites by calculating the appropriate activation energies. The analysis of Chapter 4 will allow us to incorporate exactly the effects of a free surface, and one of the aims of this chapter will be to assess the accuracy of commonly used approximations that are based on infinite-body results.

The energy change W on introducing a dislocation loop into a background stress field σ_{ij}^B may be calculated in precisely the same way as was detailed in Section 3.1. Let S be the surface, bounded by the loop, across which the displacement is allowed to be discontinuous. If σ_{ij}^D is the stress field of the dislocation loop in the absence of a background stress field σ_{ij}^B, then the energy change W is given by (3.4), with $\mathbf{b}$ the Burgers vector of the loop, and $\mathbf{m}$ the associated normal to the surface S:

$$W = -\frac{1}{2}\int_S \sigma_{ij}^D m_j b_i dS + E^c - \int_S \sigma_{ij}^B m_j b_i dS. \tag{5.2}$$

Glide loops only will be considered, so for simplicity the surface S is taken to be the region of the glide plane that is enclosed by the loop. As before, the integrand in the first integral is singular at the dislocation, and so this integral is understood to be evaluated only up to a distance q from the dislocation. If the loop expands with fixed shape, then W may be written as a function of some characteristic dimension of the loop, ρ, say. Then the activation energy

W_{act} is determined by the condition

$$\frac{dW}{d\rho} = 0. \tag{5.3}$$

The quantity

$$-\frac{1}{2}\int_S \sigma_{ij}^D m_j b_i dS + E^c \tag{5.4}$$

is called the self-energy of the loop, and will henceforth be denoted by W_s. Accurate estimation of W_s, taking the effects of a free surface rigorously into account, forms the subject matter of Section 5.2. Exact calculations for a dislocation loop of any shape are facilitated using the results of Chapter 4. The quantity

$$-\int_S \sigma_{ij}^B m_j b_i dS \tag{5.5}$$

is called the interaction energy and is henceforth denoted W_i. This term represents the energy change due to interaction between the dislocation and the background stress field when the loop is formed, which drives the nucleation process. In the absence of a background stress field σ_{ij}^B, W_i is zero, and the loop energy, (5.2), which may be written in terms of the above notation as

$$W = W_s + W_i,$$

increases monotonically with loop size, and therefore no nucleation occurs. In cases of interest, W_i is negative, and, for sufficiently large loops, renders W negative, making nucleation an energetically favoured process.

A key component in the explanation to be proferred of the experimental observations will be the distinction between two modes of nucleation, homogeneous and heterogeneous, which are now described. The stress field in a lattice mismatched layer consists of two parts. There is a uniform part with components given by (2.9), which is precisely the stress field that would exist if the idealization were realized of an atomically flat interface, and a layer free from defects and impurities. The additional part arises due to divergence from this idealization, and comprises a collection of localized stress fluctuations due to impurities, defects, interface ledges, inhomogeneities in the alloy, and so on. If the first, uniform, stress component is enough to drive nucleation on an observable timescale, then nucleation may occur at all points across the layer. Such nucleation is called homogeneous. However, if the uniform component of the stress cannot drive an observable nucleation rate than nucleation can only occur at sites of high local stress, which are described by the fluctuations in the second part of the stress field. Such nucleation is called heterogeneous.

In Section 5.3 we calcuate the activation energy for nucleation in a mismatch-induced stress

field of the type given by (2.9) with lattice mismatch given by (3.9). Using (5.1) it is then possible to determine the value of the fractional lattice mismatch, f_m, at which homogeneous nucleation is possible over a practically relevant timescale. Rates obtained using approximations based on infinite-body results are found to be seriously in error. Using these results, the observed increase in threading segment density and decrease in dislocation length as f_m nears 0.02 may be explained by postulating a transfer from heterogeneous to homogeneous nucleation. The change in observed dislocation type from 60° to 90° is explained in Section 5.4 through the description and modelling of a mechanism by which 60° dislocations at the interface are converted into 90° dislocations via a nucleation event at the site of the 60° dislocation. In this way a moderately complete description of dislocation distributions for $f_m < 0.035$ is given.

To conclude the chapter, Section 5.5 is devoted to some closing remarks concerning the significance of the pre-exponential factor in (5.1) and of energies associated with the dislocation core, which cannot be modelled using the linear theory of elasticity.

5.2 The Self-Energy of a Dislocation Loop in a Half Space

The self energies of a dislocation loop buried in a half space and of a dislocation half loop at a free surface are calculated rigorously in this section and compared with commonly used approximations for these energies, which are based on infinite-body results. Using the line-integral representation of Section 4.1 or Section 4.4 for the dislocation stresses, the exact elastic self-energy of a loop is given by

$$W_s = -\frac{1}{2}\int_S b_p dS_q \oint_C b_s \left(S^{\infty}_{pqrs}(\mathbf{x},\mathbf{x}') + S^{I}_{pqrs}(\mathbf{x},\mathbf{x}')\right) dx'_r + E^c \tag{5.6}$$

when the loop is buried in a half-space, and

$$W_s = -\frac{1}{2}\int_S b_p dS_q \oint_C b_s \left(S^{\infty\star}_{pqrs}(\mathbf{x},\mathbf{x}') + S^{I\star}_{pqrs}(\mathbf{x},\mathbf{x}')\right) dx'_r + E^c \tag{5.7}$$

when a surface half-loop is considered. The two commonly applied approximations are:

(a) for a loop buried in a half-space the energy is taken to be that of the corresponding loop in an infinite body;

(b) for a surface half-loop the energy is taken to be half that of the completed loop in an infinite body.

In Chapter 4 the line-integral representations have been constructed so that integration of the first term in the integrand of (5.6) yields the value taken in approximation (a), and integration of the first term in the integrand of (5.7) yields the value taken in approximation (b). Therefore, the corrections that must be applied to the energies given by approximations (a) and (b) are, respectively,

$$W^I = -\frac{1}{2}\int_S b_p dS_q \oint_C b_s S^I_{pqrs}(\mathbf{x},\mathbf{x}')dx'_r \tag{5.8}$$

and

$$W^{I\star} = -\frac{1}{2}\int_S b_p dS_q \oint_C b_s S^{I\star}_{pqrs}(\mathbf{x},\mathbf{x}')dx'_r. \tag{5.9}$$

Before proceeding with evaluation of these correction terms we present some known infinite-body results that allow evaluation of the approximate energies, as prescribed by (a) and (b) above, for the two geometries we shall consider: circular and rectangular loops. The self-energy of a circular glide loop in an infinite body is given by[4]

$$W_s^\infty = \frac{\mu b^2 r}{4}\left(\frac{2-\nu}{1-\nu}\right)\left[\ln\left(\frac{8r}{q}\right)-2\right] - \frac{\mu b^2 r(1-2\nu)}{16(1-\nu)^2}, \tag{5.10}$$

where r is the radius of the loop, and $b = |\mathbf{b}|$. The core-traction contribution, E^c, has been included in (5.10). Now consider a rectangular glide loop, with dimensions Y and Z, in an infinite body. Let $\mathbf{l}$ be a unit vector along one of the sides of length Z. Then defining

$$b_s^2 = (\mathbf{b}\cdot\mathbf{l})^2$$

and

$$b_e^2 = b^2 - b_s^2,$$

the self-energy of the loop is given by[24][1]

$$\begin{aligned} W_s^\infty \;=\; & \frac{\mu}{2\pi(1-\nu)}\left\{[b_e^2+(1-\nu)b_s^2]\left[Z\ln\left(\frac{2YZ}{q(D+Z)}\right)+D-Y-Z\right]\right. \\ & \left.+\,[(1-\nu)b_e^2+b_s^2]\left[Y\ln\left(\frac{2YZ}{q(D+Y)}\right)+D-Y-Z\right]\right\} \\ & +\frac{\mu(2\nu-1)}{8\pi(1-\nu)^2}(Zb_e^2+Yb_s^2). \end{aligned} \tag{5.11}$$

[1] Equation (8) of [24] contains an error discovered after publication: the term in that equation that is multiplied by $\nu b_e b_s$ should not appear. However, the results of that paper are unaffected, since this erroneous term in any case disappears for a square loop, which was the only case considered quantitatively there.

Again, the term E^c has been included. The quantites b_s and b_e are, respectively, the strengths of the screw and edge components of the sides of length Z. Equations (5.10) and (5.11) allow the approximate energies, prescribed by (a) and (b) above, to be calculated for circular and rectangular buried loops and surface half loops.

We now return to the correction terms of (5.8) and (5.9), which must be added to the approximate energies described by (a) and (b) above in order to obtain the correct half-space energies. Elementary dimensional analysis allows us to present the correction terms W^I and $W^{I\star}$ in a conveniently general form. First consider a loop buried in a half-space, and let h be the maximum distance from the free surface of a point on the loop. (If the loop is considered to have formed at the interface between a strained layer and its substrate, then h is the thickness of the strained layer.) Further, let ρ be some characteristic dimension of the loop, say the length of dislocation line. Fixing the *shape* of the loop and keeping ρ/h constant, consider varying ρ. Since the image stresses are linear in the modulus of the Burgers vector $b = |\mathbf{b}|$, dimensional analysis reveals that their values at non-dimensionalized coordinates are proportional to b/ρ. Since the area of the loop surface S is proportional to ρ^2, W^I must be proportional to ρ if b, ρ/h, and the loop shape remain fixed. Therefore we may write

$$W^I = \mu b^2 \rho \Theta(\rho/h), \tag{5.12}$$

where the form of the functional dependence, upon ρ/h, of the dimensionless correction coefficient Θ depends upon the loop shape, its orientation relative to the free surface, and the orientation of the Burgers vector.

When considering $W^{I\star}$ the situation is somewhat simplified, because for a surface half-loop of fixed shape there is only one characteristic length ρ, which we again take to be the length of dislocation line. In this case the energy correction may be written

$$W^{I\star} = \mu b^2 \rho \Theta^\star, \tag{5.13}$$

where $\Theta^\star$ is a dimensionless correction coefficient, which depends upon the loop shape, its orientation relative to the free surface, and the orientation of the Burgers vector.

It should be noted that an equation of the form of (5.12) or (5.13) cannot hold for the *total* loop energy, because of the singular part of the dislocation stress field: in this case the surface integral involves a further characteristic length, say q, which is the core cutoff parameter. This third length appears because the energy integral for the total stress field cannot be evaluated all the way up to the dislocation line where the stresses are singular. In this case the energy

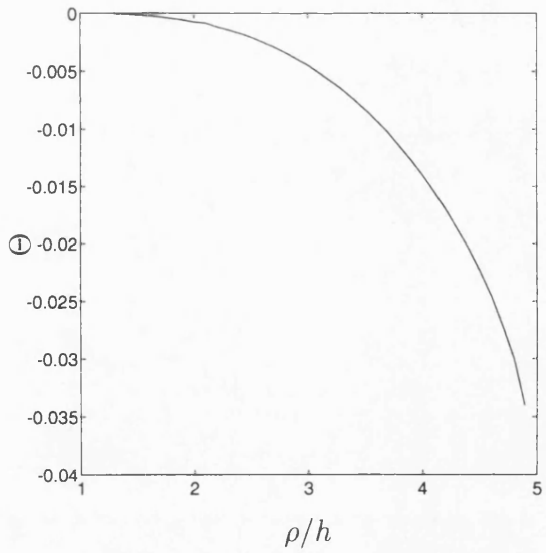

Figure 5.2: Correction coefficient Θ for a square loop buried in a half space, plotted as a function of the characteristic ratio ρ/h; ρ is the length of dislocation line and h is the maximum distance from a point on the loop to the free surface. The loop was taken to be formed on a (111) plane in a half-space with (100) surface normal and to have $\frac{1}{2}[10\bar{1}]$ Burgers vector.

consists not only of terms proportional to ρ but also terms linear in $\ln(\rho/q)$.

In presenting results for the correction terms of (5.12) and (5.13) we restrict our attention to the formation of a pure glide loop with $\frac{1}{2}[10\bar{1}]$ Burgers vector emanating from a (100) surface on the (111) glide plane. Such a loop would, in a (100) strained-layer structure, form at the interface a dislocation of the commonly observed 60° type. In all calculations the value of Poisson's ratio was taken to be 0.25. In fact the dependence of the results on Poisson's ratio was found to be weak, with the correction terms varying by less than 7% from the values given here as Poisson's ratio was varied between 0.2 and 0.4. Having fixed the orientation of the loop and its Burgers vector, the function $\Theta(\rho/h)$ and the number Θ^* depend only upon the loop *shape* considered.

First consider formation of a square glide-loop, buried in a half-space. The infinite-body energy of a square glide-loop is obtained from (5.11) by putting $Y = Z$. Therefore, we need only consider explicitly the correction term W^I. Since a square is assembled from straight dislocation segments, only the surface integral required in (5.8) need be assessed numerically, the line integral having been evaluated analytically in Section 4.3. The function $\Theta(\rho/h)$ for a square loop is plotted in Fig. 5.2. Note that since the (111) glide plane of the loop is inclined at an angle $\tan^{-1}(1/\sqrt{2})$ to the surface normal, (100), and since each side of the square is of length $\rho/4$, the condition

$$\rho/h = 2\sqrt{6} \tag{5.14}$$

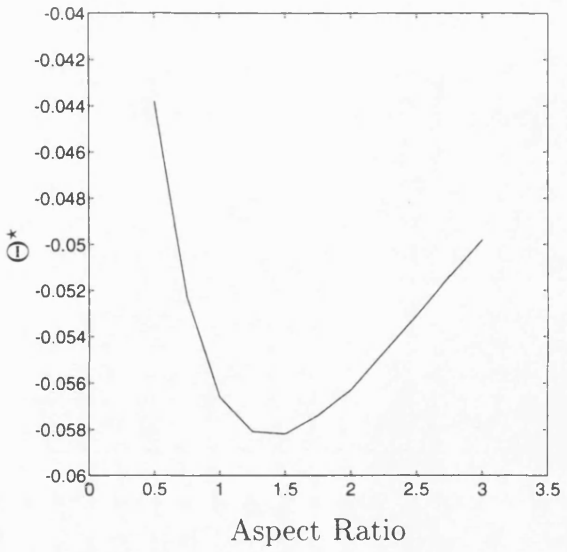

Figure 5.3: Correction coefficent $\Theta^\star$ for a rectangular surface half-loop plotted as a function of aspect ratio of the loop, which is defined to be the length of the side parallel to the surface divided by the length of one of the sides that meets the free surface. The loop was taken to be formed on a (111) plane at a (100) surface and to have $\frac{1}{2}[10\bar{1}]$ Burgers vector.

determines the point at which the loop meets the free surface. At this value of ρ/h, Θ has a negative logarithmic singularity, for the reasons discussed in Section 4.4. It can be seen that the error incurred in approximation (a) increases rapidly once ρ/h exceeds the value three, corresponding to the loop having expanded approximately two-thirds of the way from the interface of a strained layer to its surface.

For the case of a surface half-loop we consider two shapes: a half-rectangle and a semicircle. For these geometries the infinite-body results required by approximation (b) may be obtained from (5.10) and (5.11). Again, for a rectangular half-loop, which may be built-up from straight dislocation segments, only the surface integral appearing in (5.9) need be evaluated numerically. However, for a semicircular surface half-loop, the surface and line integrals must both be evaluated numerically. In Fig. 5.3 the shape factor $\Theta^\star$, for a rectangular half-loop, is plotted as a function of the aspect ratio of the loop. The aspect ratio has been defined to be the length of the side parallel to the surface divided by the length of one of the sides that meets the free surface.

For a semicircular surface half-loop we obtain the value

$$\Theta^\star = -0.034.$$

5.3 Homogeneous Nucleation

The analysis of the previous section enables calculation of the self-energy, W_s, of rectangular and circular dislocation loops, as given by (5.4), taking fully into account the effects of the free surface. To obtain the activation energy, the sum of W_s and W_i must be maximized, where W_i is the energy change due to interaction between the dislocation and the background stress field σ_{ij}^B. In this section we restrict attention to homogeneous nucleation, and so take σ_{ij}^B to consist solely of the stresses due to a uniform lattice mismatch, e_{ij}^m, between the lattice constants of the layer and substrate, as given by (2.9). As in previous sections we shall consider the case of cubic materials, which is the case applicable to semiconductor materials, so that the lattice mismatch is given by (3.9). For the case of isotropy the non-zero stress components in the layer then become:

$$\sigma_{11}^B = \sigma_{33}^B = -2\mu \left(\frac{1+\nu}{1-\nu} \right) f_m. \tag{5.15}$$

Here we have reverted to a coordinate system which has the x_2-axis normal to the free surface. For the stresses given in (5.15) the integral required by (5.5) is trivial. We obtain

$$W_i = 2\mu \left(\frac{1+\nu}{1-\nu} \right) f_m b_1 A \cos\alpha, \tag{5.16}$$

where A is the area of the plane surface bounded by the loop, and, as before, α is the angle between the glide plane and the normal to the free surface. The glide plane has been taken to lie parallel to the x_3-axis, so that b_1 is the component of the Burgers vector that is parallel to the free surface and perpendicular to the line where the glide plane meets the free surface. For a misfit dislocation f_m and b_1 are of opposite sign so that W_i is negative. With W_i given by (5.16) and W_s as calculated in the previous section we can calculate the activation energies, using (5.3), and nucleation rates, using (5.1), for loop formation in a layer stressed due to a fractional lattice mismatch f_m.

The rates that we present are all for nucleation in an $In_xGa_{1-x}As/GaAs(001)$ strained-layer system, and we take $\mu = 4.87 \times 10^{10}$Pa and $\nu = 0.25$. The fractional lattice mismatch is then related to the In fraction x by (3.10). The activation energy is obtained for the loop shapes considered by maximizing the sum of the self-energy, W_s (obtained using the results of the previous section), and the energy reduction due to strain relief (which is given in (5.16)). First consider nucleation of a square loop buried in a strained layer. The centre of the loop is at a depth y below the free surface. The nucleation rate, obtained via (5.1), is plotted as a function of y for layers containing In fractions x =0.1, 0.3, and 0.4 in Figs. 5.4, 5.5, and 5.6, respectively;

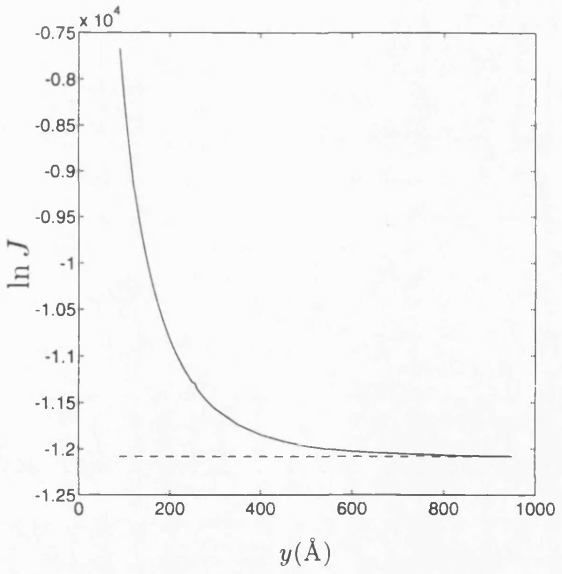

Figure 5.4: Natural logarithm of the nucleation rate versus depth of loop below the surface for a square loop buried in an $In_{0.1}Ga_{0.9}As$ (001) oriented strained layer. The dislocation was taken to be formed on a $\{111\}$ plane and to have $\frac{1}{2}[10\bar{1}]$ Burgers vector. The solid line gives the rates obtained using exact half-space energies, and the dashed line gives those obtained using the approximate energies. The rates were put into units of cc^{-1} before taking the logarithm.

i.e. for $f_m = 0.0071$, 0.021 and 0.028. The temperature was taken to be 800K in all cases, corresponding to typical growth and annealing temperatures of 500-550°C. The solid curves were obtained using the exact expression for the loop self-energy, and the dashed curves were obtained using approximation (a), which takes the loop self-energy just to be the infinite-body value. Hence the values given by the dashed lines do not vary with y. For all values of x the rates obtained using the exact half-space energies exceed those obtained using the approximate energies by very many orders of magnitude when y is small. However, it should be noted that the active region of a device may have a volume as small as approximately 10^{-10}cc. In such a volume, a rate of $J = 1cc^{-1}s^{-1}$ corresponds to one nucleation event in the prescribed volume every 10^{10} seconds, i.e. every 3000 years! Therefore, a nucleation process for which $\ln J < 0$ may certainly be discarded as irrelevant at the temperature for which the rate was calculated. We require an 'observability condition' by which to judge whether a rate is experimentally significant. An experimentally significant rate might be taken to be one nucleation event in the volume of 10^{-10}cc every minute, so that dislocations will be introduced in numbers over an experimental timescale. This corresponds to a rate of the order of $10^8 cc^{-1}s^{-1}$, for which $\ln J \approx 20$. Therefore, when the In fraction is as high as 0.4, nucleation within about 20Å of the free surface, driven solely by the mismatch-induced stress field, may be significant. However, nucleation so close to

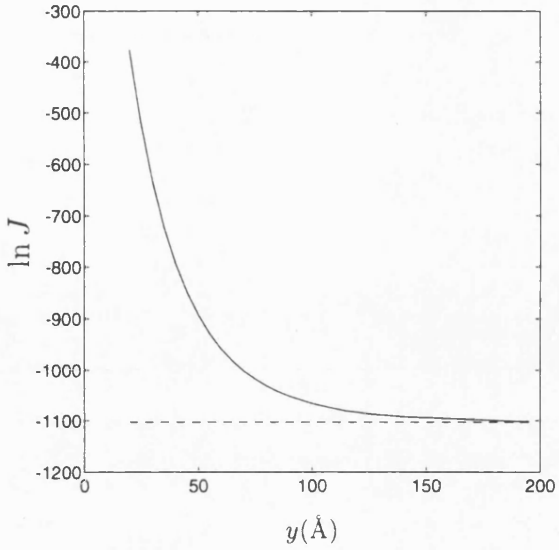

Figure 5.5: As for Fig. 5.4, but for an In fraction of $x = 0.3$.

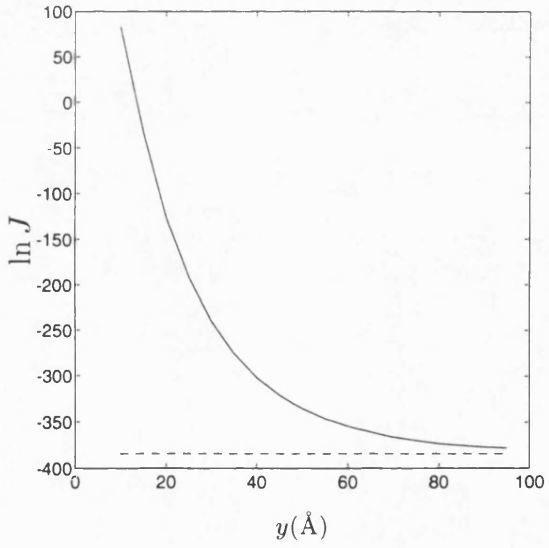

Figure 5.6: As for Fig. 5.4, but for an In fraction of $x = 0.4$.

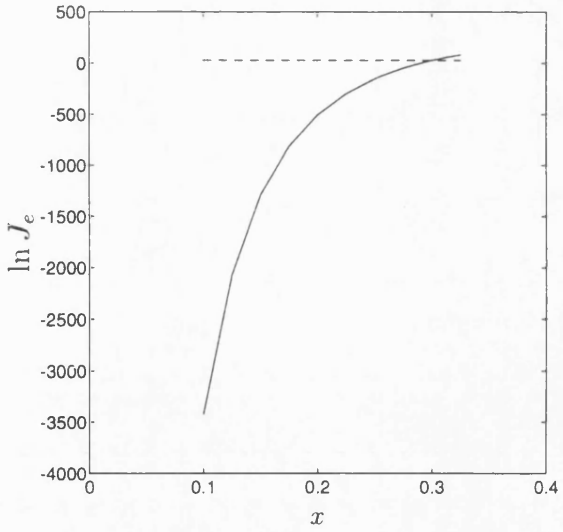

Figure 5.7: Natural logarithm of the nucleation rate for a surface half square, obtained using the exact elastic self-energy, plotted as a function of In fraction x. The broken line marks the observability criterion, $\ln J \approx 30$. The nucleation rate was put into units of $\mathrm{cc}^{-1}\mathrm{s}^{-1}$ before taking the logarithm.

the surface is difficult to distinguish from actual half-loop nucleation, since the loop can reach the surface before attaining its critical size. The conclusion to be obtained from these results must be, therefore, that homogeneous nucleation within the layer is very unlikely for the values of f_m considered here. These conclusions also hold for the $\mathrm{Ge}_x\mathrm{Si}_{1-x}/\mathrm{Si}$ system.

Results for nucleation of surface half-loops are plotted in Figs. 5.7, 5.8, and 5.9. We have considered nucleation of a surface half loop for different values of In fraction x. In all cases the loop was taken to have formed on a $\{111\}$ plane in a (001) oriented structure, and to have Burgers vector $\frac{1}{2}[10\bar{1}]$. In the following, J_e refers to the nucleation rate obtained, via (5.1), using the exact formulation for the half-loop, and J_a refers to the rate obtained using approximation (b) (i.e. the half-loop energy is taken to be half that of the completed loop in a whole space). In Fig. 5.7 the logarithm of the nucleation rate J_e is plotted as a function of In fraction x. The nucleation rate given in Fig. 5.7 is only relevant to nucleation at the free surface. Quite what is meant by 'at the free surface' is difficult to make precise, but it seems not unreasonable to consider nucleation within 10Å of the free surface to be determined by the half-loop activation energy. The rate at which dislocations are created by half loop nucleation is then given, per unit surface area of the device, by

$$R = 10^{-7} J_e \mathrm{cm}^{-2}\mathrm{s}^{-1},$$

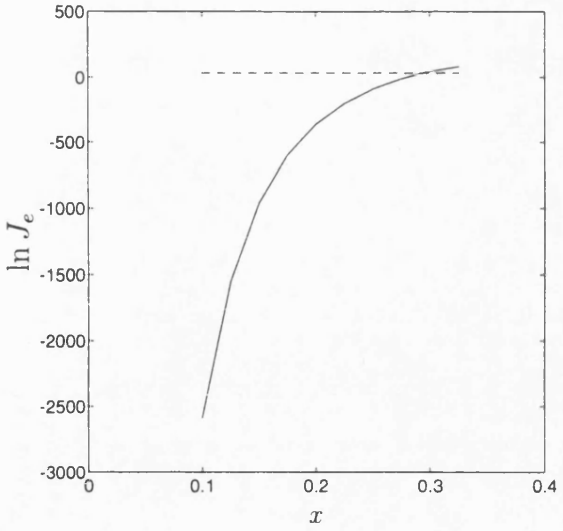

Figure 5.8: As for Fig. 5.7 but for a temperature of 1050K.

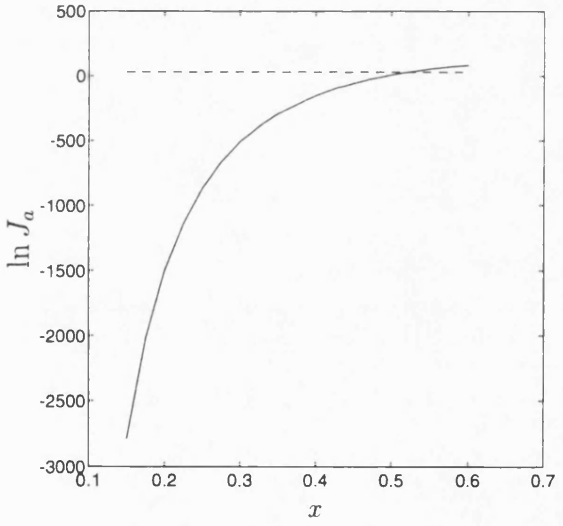

Figure 5.9: Natural logarithm of the nucleation rate for a surface half square, obtained using the approximate elastic self-energy, plotted as a function of In fraction x. The broken line marks the observability criterion, $\ln J \approx 30$. The nucleation rate was put into units of $\mathrm{cc}^{-1}\mathrm{s}^{-1}$ before taking the logarithm.

where the rate J_e is in units of $cc^{-1}s^{-1}$. The particular choice of this distance over which nucleation is considered to be, effectively, at the surface affects only very weakly conclusions about conditions under which dislocation nucleation is likely to be significant. This is because the rate J_e is so strongly dependent upon the activation energy in the Arrhenius exponential. A rate of the order of $10^6 cm^{-2}s^{-1}$ corresponds to a nucleation event every minute or so in each $1\mu m \times 1\mu m$ area of the device surface. Such a rate is certainly observable, and corresponds to $\ln J \approx 30$, which we take to be our observability criterion for half loop nucleation. For a surface half square this value, as marked by the broken line in Fig. 5.7, is attained when $x = 0.3$, or equivalently, from (3.10), $f_m = 0.021$. Notice that small variations of ± 0.04 from this value of x yield rates of the order of $10^{28} cm^{-2}s^{-1}$ and $10^{-72} cm^{-2}s^{-1}$, which are, respectively, huge and tiny. This reveals that the transition from surface nucleation being virtually non-existant to surface nucleation occuring at will occurs over a very small range of x. Therefore, the particular observability condition chosen does not strongly effect the predicted threshold lattice mismatch at which homogeneous nucleation at the free surface becomes 'switched on'. The effect of variation in temperature is illustrated in Fig. 5.8 where the rate J_e is plotted as a function of x for a temperature of 1050K. Although the increase in temperature causes a large increase in the nucleation rate for a given x, the threshold value of x at which the observability criterion is satisfied is little affected, as is evident from the figure. Also, we note here that the threshold mismatch in the Ge_xSi_{1-x}/Si system is almost identical with the value of $f_m = 0.021$ obtained here for the $In_xGa_{1-x}As/GaAs$ system and that the results for surface semicircles are very similar to those presented here for surface half squares. We conclude that in both material systems, homogeneous nucleation of dislocations at the free surface becomes significant as f_m exceeds 0.02.

The logarithm of the rate J_a, obtained using the approximate elastic self-energy, is given in Fig. 5.9. Note that, for a given x, J_a is very many orders of magnitude less than J_e. Moreover, note that the observability condition $\ln J_a = 30$ is not attained until $x = 0.52$, which gives a very much larger value for threshold mismatch than that obtained using the exact elastic self-energy: specifically, using this approximation homogeneous nucleation is not expected until f_m=0.037.

It may be concluded from the results obtained above that homogeneous nucleation should be quite insignificant if f_m is less than 0.02. For higher values of f_m, homogeneous nucleation of surface half loops should begin to occur at a rapid rate. Homogeneous nucleation of loops buried within the layer should not be significant for the range of f_m considered here. Therefore, for $f_m < 0.02$, dislocations may only be formed through heterogeneous nucleation at regions of local high stress. As f_m increases through 0.02, homogeneous nucleation will 'switch on'.

The effect that this change in the nucleation mode will have on the structure of the dislocation distribution is now discussed.

When nucleation is heterogeneous, as it is for $f_m < 0.02$, the number of nucleation sites is obviously limited by the number of local inhomogeneities in the stress field. In the high quality layers that may be grown today, this number is not expected to be large. When a dislocation forms at an inhomogeneity, its stress field inhibits the nucleation of further dislocations at the inhomogeneity. If the inhomogeneity in the stress field is very large, then several dislocations may be nucleated before the stress fields of the nucleated dislocations 'deactivate' the inhomogeneity as a nucleation site. In this way clustering of dislocations may arise. Clustering may also arise through dislocation multiplication. Now, the threading segment density is obviously a direct measure of the number of dislocations that have nucleated. If the strained layer is of reasonably high quality, then the number of inhomogeneities will be small, and hence the number of dislocations and threading segments will be small also. Because the number of dislocations that may be nucleated is small, the strain must be relaxed by extension of the dislocations that have successfully nucleated, rather than by the nucleation of large numbers of new dislocations, and so long misfit segments are to be expected. It is clear, then, that the restriction of nucleation to heterogeneous mode can explain the experimental observations of low threading segment density, long misfit segment length, and dislocation clustering, at low values of f_m. Precise quantitative modelling of heterogeneous nucleation is made difficult by the range of possible nucleation sites. Steps towards the modelling of nucleation at non-planar interface features are taken in Chapter 6, and nucleation at a dislocation is considered later in this chapter. However, the qualitative explanation given above will apply regardless of the particular nature of the heterogeneous nucleation site.

When f_m reaches 0.02, homogeneous nucleation at the free surface becomes possible over an experimentally significant timescale, and dislocations may form at any point across the layer. This immediately explains the shift from a clustered to a random distribution. The manner in which the onset of homogeneous nucleation affects the dislocation distribution depends on the relative timescales of nucleation and extension of dislocations. If the timescale for nucleation is long compared with that for extension, then the strain in the system will be relaxed by extension of relatively few dislocations in the layer rather than by nucleation of large numbers of further dislocations. In this case, long misfit segments should be observed. By contrast, if the timescale for nucleation is short compared with that for extension, then the strain will be relaxed via the nucleation of large numbers of dislocations, rather than by extension to long lengths of a few dislocations. The dislocations observed will, correspondingly, be rather short. It is possible to

illustrate and quantify this argument in a rudimentary way.

Consider a strained layer with dislocation loops nucleating homogeneously at its surface. Because of the stress field of each loop, the activation energy for nucleation of further surface half loops will be raised in the vicinity of that loop. As a result of the strong exponential dependence of the rate upon the activation energy, there will be an area of surface around each loop (which will increase as the loop expands), over which further nucleation will be effectively halted, and so the total nucleation rate will be reduced from its initial value as the loops expand. Again as a result of this strong exponential dependence, it may be anticipated that the region of the surface over which the nucleation rate is reduced but still observable will be rather small, and may be ignored. Hence, we treat the surface as being divided into regions over which nucleation is virtually halted and regions over which the nucleation rate has its initial value. Let $A(t)$ be the area of the free surface, at time t, over which nucleation is *unaffected* by the presence of other dislocations, and let the nucleation rate (per unit area) over this area be R, which is assumed constant. Then the time rate of change of the total number of dislocations, N, is given by

$$\frac{dN}{dt} = RA(t).$$

Suppose that each dislocation as it grows reduces the area available for nucleation at a constant rate α. (This is a very crude approximation, made purely to enable an order of magnitude estimate). Then

$$\frac{dA}{dt} = -\alpha N.$$

Imposing the initial conditions

$$N(0) = 0, \quad \frac{dN}{dt}(0) = RA_0,$$

it is trivial to show that the above system of equations yields the solution

$$N(t) = \sqrt{\frac{R}{\alpha}} A_0 \sin(t\sqrt{\alpha R}).$$

Nucleation stops at time $t = \pi/(2\sqrt{\alpha R})$, since R and A are never negative. Assuming that the nucleation process runs its course, the final areal dislocation density is given by

$$\frac{N}{A_0} = \sqrt{\frac{R}{\alpha}}.$$

(The threading segment density is twice this, since each nucleation event produces two threading

segments.) Therefore, the final dislocation concentration depends on the ratio of the timescales of dislocation nucleation and expansion. If dislocations expand very fast, then nucleation will be prevented over a wide area before many dislocations have formed. If expansion is slow, then many dislocations can form before the available nucleation area is significantly reduced.

We now require an estimate for α, so that we can obtain an order of magnitude estimate for the rate required to produce a significant number of dislocations. Jain, Jain, Harker, and Bullough[40], who have recently considered heterogeneous and homogeneous nucleation processes, found that a loop only has a significant effect on the activation energy of a neighbouring loop over a range approximately equal to the diameter of the loop. Therefore, nucleation will be inhibited over an area of the same order as the area of the loop. Expansion of the loop is constrained, in one direction, by the interface. Therefore, once the loop has reached the interface, its area increases only by increase of the length of the side parallel to the inteface. A typical layer thickness is of the order of 0.1μm and a typical dislocation velocity (i.e. rate of expansion of a misfit segment in the interface) is 1-100μms^{-1}[66], depending on the lattice mismatch. A reasonable choice for α, for the purposes of obtaining an order of magnitude estimate, is therefore $\alpha \approx 10^{-8}\text{cm}^{-2}\text{s}^{-1}$. Now recall that the nucleation rate per unit surface area is related to the volume nuleation rate by the approximate relationship $R = 10^{-7}J$, where the units of R are $\text{cm}^{-2}\text{s}^{-1}$ and those of J are $\text{cc}^{-1}\text{s}^{-1}$. Use of this relationship in conjunction with (5.1) and the above equation yields

$$\frac{N_{thread}}{A_0} \approx 10^{19} \exp(-W_{act}/2kT)\ \text{cm}^{-2},$$

which is the areal threading dislocation density due to *homogeneous* nucleation. This density is plotted as a function of mismatch strain f_m in Fig. 5.10 for the $\text{Ge}_x\text{Si}_{1-x}$/GeSi(100) system, where some experimental values obtained by Kvam em et al.[44] are included for comparison. The solid line gives the threading segment density due to homogeneous nucleation using the pre-exponential factor of 10^{36} in (5.1). The dashed line was obtained using a pre-exponential factor of 10^{20} in (5.1), instead of the factor 10^{36}. This leads to the factor of 10^{19} in the above equation being changed to 10^{11}. Preliminary development of a model of loop nucleation based on kink formation has suggested a pre-exponential of the order of 10^{20}. This model is discussed further in Section 5.5. In addition to the threading dislocations arising through homogeneous nucleation, an additional number will arise from heterogeneous nucleation sites. At low mismatch, all nucleation is heterogeneous, and the threading segment density was found by Kvam *et al.*[44] to be 10^8cm^{-2}. The number of dislocations that are nucleated heterogeneously is governed by the number of stress inhomogeneities, which is not expected to be strongly dependent on f_m. We therefore assume that the threading segment density due to heterogeneous nucleation

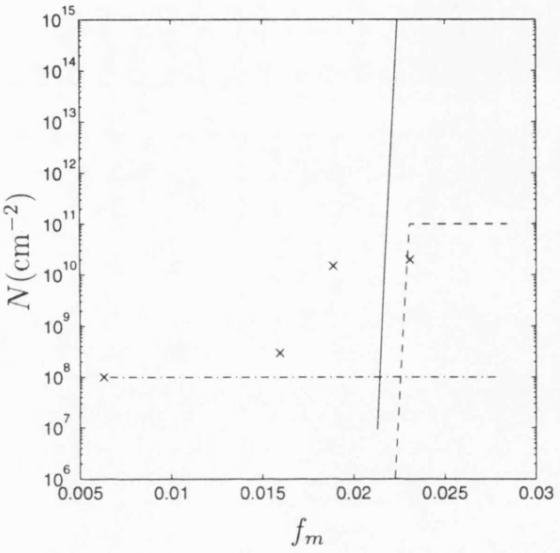

Figure 5.10: Threading segment density plotted as a function of mismatch. The solid and dashed lines give the predicted densities arising because of homogeneous nucleation using pre-exponential factors of, respectively, 10^{36} and 10^{20} in (5.1). The dot-dashed line gives the density arising through heterogeneous nucleation, assuming this density to be constant at its low-mismatch value. The total threading segment density is the sum of the densities arising from homogeneous and heterogeneous nucleation. The crosses mark total threading segment densities obtained experimentally.

is approximately constant at the value of $10^8 \mathrm{cm}^{-2}$. The dot-dashed line in Fig. 5.10 marks this density. The total threading segment density is given by the sum of the densities due to heterogenous and homogeneous nucleation. For both choices of pre-exponential factor a sharp increase in the threading segment density is predicted to occur as f_m increases between 0.02 and 0.025. This is in good agreement with the position of the observed increase, which occurs for values of f_m between 0.015 and 0.02. However, when the pre-exponential factor of 10^{36} is used then the threading segment density is predicted to plateau at a value of around $10^{19}\mathrm{cm}^{-2}$, which is very much higher than the experimentally observed value of $2 \times 10^{10}\mathrm{cm}^{-2}$. When the pre-exponential factor of 10^{20} is used the threading segment density is predicted to plateau at $10^{11}\mathrm{cm}^{-2}$, which is in much better agreement with experiment. The observed shortening of the dislocations as the mismatch is increased through the range 0.015 to 0.025 follows immediately from the increased dislocation density, since the strain relieved by a collection of dislocations is proportional to the density of dislocations multiplied by their mean length: since the dislocation density increases by two orders of magnitude over a rather small range of mismatch strain, the mean length of the dislocations must fall accordingly. Moreover, the total time taken for the nucleation stage, $t \approx \pi/(2\sqrt{\alpha R})$ becomes so short (less than a second when homogeneous nucleation becomes significant) that all of the dislocations nucleate nearly simultaneously and so will all be close to the mean length.

Note that the predicted rise in threading segment density is more rapid than that observed experimentally. This is easily explained. The theoretical results are obtained on the assumption that the mismatch strain is uniform throughout the layer. Therefore, all parts of the layer are predicted to exceed the threshold mismatch simultaneously. In practice, the composition of the alloy will not be uniform, and so the lattice mismatch will fluctuate around a mean value of f_m. For this reason different parts of the layer pass through the threshold mismatch at different stages as the mean mismatch f_m is increased, resulting in a more gradual change in the dislocation distribution than is suggested by the theoretical results above.

Only the nucleation of loops that lead to formation of 60° dislocations at the interface has been considered here. However, an interesting feature of systems in which the mismatch strain is greater than about 0.02 is that the dislocations observed are of the 90° type. These run in the $\langle 110 \rangle$ interface directions and have Burgers vector along the interfacial $\langle 110 \rangle/2$ lattice vector that is perpendicular to the dislocation line. The formation of such dislocations, and the relationship between their rate of nucleation and the nucleation rates given above for 60° dislocations, form the subject matter of the next section.

5.4 The Conversion of 60° into 90° Dislocations

In Section 5.3 we considered the nucleation of a new dislocation in the presence of a stress field due solely to the lattice mismatch in a structure with a planar interface. In this way we provided conditions under which homogeneous nucleation is possible, and related the observed increase in threading dislocation density, decrease in dislocation length, and shift from clustered to random distribution at $f_m \approx 0.02$ to the onset of homogeneous nucleation. The calculations of the previous section were made assuming that the loops form 60° dislocations at the interface. However, it has been observed[34, 12, 44] that for both the Ge_xSi_{1-x}/Si(100) and $In_xGa_{1-x}As$/GaAs(100) strained layer systems there is a change in the dominant dislocation type from the commonly observed 60° mixed type at low mismatch strains ($f_m < \sim 0.015$) to 90° pure edge type at high mismatch strains ($f_m > \sim 0.023\%$). The changes in the dislocation distribution that were explained in the previous section are accompanied by a change in dislocation type. It is this accompanying change that we now consider. Energetically, 90° dislocations are preferable to the 60° type, but because of crystallographic constraints they are more difficult to form in practice: glide on the close-packed {111} glide planes of a 60° dislocation is preferable to that on the {100} glide planes of a 90° dislocation. Moreover, since the glide plane of a 60° dislocation is inclined to the interface, threading motion and loop formation may occur for this dislocation type by pure glide, whereas the same nucleation processes are only available for a 90° dislocation via climb in a (100)-oriented structure, since in this case the glide plane and interface coincide. In this section we describe and model a mechanism by which a 90° dislocation may be formed, through a nucleation event taking place at the site of a pre-existing 60° dislocation. The nucleation rate obtained is found to attain an experimentally significant value at a mismatch strain of around 0.02, in excellent agreement with experiment. Because of the nature of the mechanism, the rate and configuration determining factor is the nucleation of 60° dislocations, as considered in Section 5.3.

Because of the crystallographic constraints on direct formation of 90° dislocations it has commonly been assumed that this dislocation type must be formed by some appropriate interaction of 60° dislocations; in fact there is also some experimental evidence for the occurence of such interactions[12, 44, 53]. A commonly suggested mechanism is the combination of two 60° dislocations, perhaps surface half-loops, with appropriately oriented Burgers vectors, to form a 90° dislocation. In terms of Burgers vectors a typical such reaction may be written

$$\frac{1}{2}[101] + \frac{1}{2}[\bar{1}10] \longrightarrow \frac{1}{2}[011]. \tag{5.17}$$

Kvam *et al.*[44] and Willis *et al.*[75] observed that a complementary 60° dislocation may also be nucleated directly at the site of a pre-existing 60° dislocation at the interface, thereby forming a 90° dislocation. In this case the complementary 60° dislocation is nucleated via formation of a loop, the upper segment of which leaves the layer at the free surface. In this case the reaction may be written

$$\frac{1}{2}[\bar{1}10] \longrightarrow \frac{1}{2}[011] + \frac{1}{2}[\bar{1}0\bar{1}]. \tag{5.18}$$

Hull and Bean[34] postulated that 90° dislocations may form via the direct nucleation of 90° surface climb loops. The elastic energy barrier to formation of these is found to be lower than for 60° loops because of the greater strain relieving capacity of a 90° dislocation. However, the mass transport required for motion of such a climb loop would seem likely to undermine the ability of such a mechanism to give rise to the large scale formation of 90° dislocations observed. Moreover, other authors[44, 53] have observed two threading segments on complementary {111} planes linking an end of a 90° dislocation to the free surface, an observation that is strongly indicative of some reaction of 60° glide dislocations. The first mechanism, described by (5.17), also seems unlikely to be a candidate for mass generation of 90° dislocations. Given a 60° dislocation at the interface, a 60° dislocation loop must nucleate at the surface on an appropriate glide plane so as to ensure collision with the pre-existing dislocation. The range of glide planes on which nucleation of a surface loop would yield a successful interaction would be limited, because the gliding loop would only be able to adjust its position by climb to a limited degree as it approached the interface. Moreover, the surface loop would have to have precisely the correct Burgers vector in order to convert the interfacial misfit dislocation into a 90° dislocation. That events should conspire in such a manner when the surface nucleation site is perhaps several thousand angstroms from the interface seems improbable, because the influence at the free surface of the interfacial dislocation will be small.

An alternative mechanism that might give rise to the interaction described by (5.17) is that postulated by Fitzgerald and Ast[20]. They suggested that a 90° dislocation could form when two complementary 60° dislocations glide out of the interface and combine either above or below the interface. For reasons similar to those discussed in relation to the combination of surface half-loops, it seems improbable that this mechanism would produce 90° dislocations in anything other than small numbers. For 90° dislocations to form by this method, pairs of dislocations with appropriate Burgers vectors must be distributed on complementary glide planes along the interface. Moreover, unless the members of a pair are close to one another, their combination would have to occur far from the interface, and energetic climb processes would be required to return the resulting 90° dislocation to its observed position, at or near the interface. Finally,

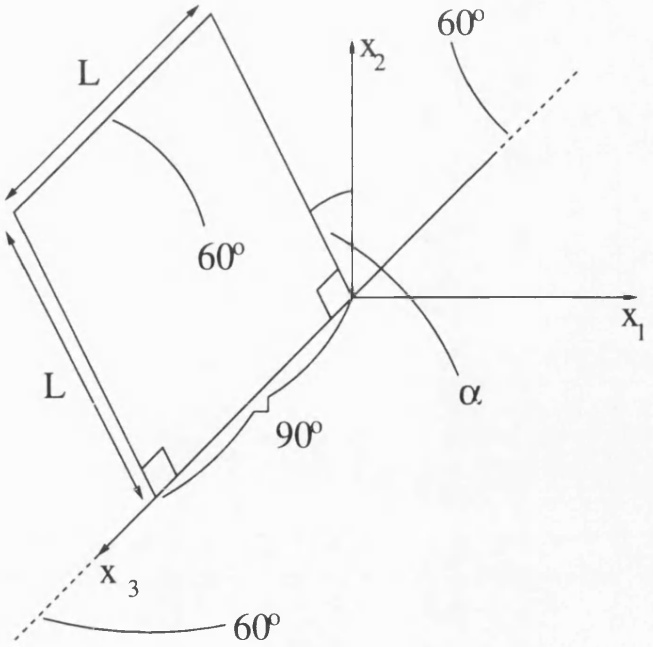

Figure 5.11: Illustration of a square dislocation loop nucleated at a pre-existing 60° dislocation. The coordinate system with respect to which Burgers vector components are referred is shown, with the pre-existing 60° dislocation chosen to lie along the x_3-axis. The interface between the strained layer and substrate is the plane $x_2 = 0$; the layer occupies some region $0 < x_2 < h$.

there is no reason for this mechanism to be mismatch dependent, and so it cannot be used to explain the observed mismatch-strain threshold for 90° dislocation formation. Therefore, the mechanism described by (5.18) seems more appealing, whereby an appropriate dislocation loop nucleates at the site of the interfacial dislocation. In this case the nucleation event is locally driven, which is intuitively satisfying, and does not require a fortuitous conspiracy of events happening perhaps thousands of angstroms apart.

In this section we examine this mechanism for the formation of 90° dislocations by calculating the elastic energy barrier to nucleation of a 60° glide loop at a complementary pre-existing 60° dislocation in a lattice-mismatched layer. By introducing the results of this calculation into (5.1) it will be illustrated that the nucleation mechanism studied here is expected to 'switch on' over a range of mismatch strains almost precisely the same as that over which 90° dislocation formation is observed to commence in experimental studies.

We consider a square dislocation loop forming at the site of a pre-existing 60° dislocation, which is modelled as being infinitely long and straight. The configuration is illustrated in Fig. 5.10. With respect to coordinates defined in Fig. 5.10 the Burgers vectors of the loop and

pre-existing misfit dislocation may be written,

$$\begin{aligned} \mathbf{b}^l &= b(-1/2, 1/\sqrt{2}, 1/2) \\ \mathbf{b}^p &= b(-1/2, -1/\sqrt{2}, -1/2) \end{aligned} \tag{5.19}$$

respectively. In defining the above Burgers vectors, the line sense of the loop has been chosen so as to coincide with that of the pre-existing dislocation along their mutual segment. Moreover, for the case of compressive strain, which is the case we shall consider, the Burgers vectors are defined by the F/S righthand screw convention relative to a positive line direction along the *negative* x3-axis. Notice that along the mutual segment of the loop and pre-existing dislocation the Burgers vector, $\mathbf{b}^l + \mathbf{b}^p = b(-1, 0, 0)$, is that of a 90° dislocation.

The self-energy of a square loop in a half space is obtained from Section 5.2. All that remains to be considered is the energy change W_i (defined by (5.5)) due to interaction with the background stress field, which consists of the mismatch stress field and the stress field of the pre-existing dislocation. The contribution to W_i from the mismatch induced stress field is given in (5.16). All that is left to evaluate is the quantity

$$-\int_S \sigma_{ij}^p m_j b_i^l dS, \tag{5.20}$$

where σ_{ij}^p is the stress field of the pre-existing 60° dislocation and S is the region of the glide plane bounded by the loop. Note that σ_{ij}^p is singular at one edge of the surface S, so the integral in (5.20) is understood to be evaluated up to a distance q from the dislocation core, and for consistency and uniqueness, a term like E^c is included to account for the interaction between the misfit dislocation and the loop over their section of mutual core boundary. For a dislocation along the x_3-axis (distance h from the free surface), with Burgers vector $\mathbf{b}$, defined relative to a positive line direction chosen in the *negative* x_3 direction, the stresses due to the *edge* component of the dislocation may be defined in terms of the Airy functions

$$\begin{aligned} \chi^\infty(x_1, x_2) &= \frac{\mu}{2\pi(1-\nu)} \mathrm{Re}[(b_2 x_1 - b_1 x_2) \ln z] \\ \chi^i(x_1, x_2) &= \frac{\mu}{2\pi(1-\nu)} \mathrm{Re}\left(2(b_1 - \mathrm{i}b_2)(x_2 - h)\frac{h}{z'} + (b_1 x_2 - b_2 x_1) \ln z'\right), \end{aligned}$$

where

$$z = x_2 + \mathrm{i}x_1,$$

$$z' = 2h - x_2 + \mathrm{i}x_1.$$

The term χ^∞ supplies the infinite-body stress field of the dislocation, and is taken from Hirth and Lothe[31], having changed the sign of their expression to account for the fact that they use the opposite Burgers vector convention to that used here. This term ensures that the discontinuity in the dislocation displacement field is correctly accounted for. The non-singular 'image' term, χ^i, ensures that the zero-traction condition is fulfilled at the free surface; in terms of stress functions, this condition becomes[15]

$$\frac{\partial}{\partial x_2}(\chi^\infty + \chi^i) = 0 \quad \text{on } x_2 = h.$$

The stresses due to the edge component of the dislocation are obtained from the above stress functions according to

$$\sigma_{11} = \frac{\partial^2 \chi}{\partial x_2^2}, \quad \sigma_{22} = \frac{\partial^2 \chi}{\partial x_1^2}$$

$$\sigma_{12} = -\frac{\partial^2 \chi}{\partial x_1 \partial x_2}, \quad \sigma_{33} = \nu(\sigma_{11} + \sigma_{22}),$$

where superscripts have been suppressed. The stresses due to the screw component of the dislocation may be written directly as[31]:

$$\sigma_{13}^\infty = -\frac{\mu b_3}{2\pi}\mathrm{Re}\left(\frac{1}{z}\right), \quad \sigma_{23}^\infty = \frac{\mu b_3}{2\pi}\mathrm{Re}\left(\frac{\mathrm{i}}{z}\right),$$

$$\sigma_{13}^i = -\frac{\mu b_3}{2\pi}\mathrm{Re}\left(\frac{1}{z'}\right), \quad \sigma_{23}^\infty = -\frac{\mu b_3}{2\pi}\mathrm{Re}\left(\frac{\mathrm{i}}{z'}\right).$$

After some lengthy but straightforward algebra it may be shown that

$$\begin{aligned}
-\int_S \sigma_{ij}^p m_j b_i^l dS \;=\; & \frac{\mu L}{2\pi(1-\nu)}[b_1^2 - b_2^2 - (1-\nu)b_3^2]\ln\left(\frac{L}{q}\right) + \frac{\mu L(1-2\nu)}{8\pi(1-\nu)}(b_1^2 + b_2^2)\cos 2\alpha \\
& + \frac{\mu L}{4\pi(1-\nu)}\left[[b_1^2 - b_2^2 - (1-\nu)b_3^2]\ln\left(\frac{4h^2}{L^2 + 4h^2 - 4hL\cos\alpha}\right)\right. \\
& 2(b_1^2 + b_2^2)\left(1 - \frac{(2h - L\cos\alpha)^2}{L^2\sin^2\alpha + (2h - L\cos\alpha)^2}\right) \\
& + (b_2^2 - b_1^2)\left(1 - 4h(h - L\cos\alpha)\frac{(2h - L\cos\alpha)^2 - L^2\sin^2\alpha}{[L^2\sin^2\alpha + (2h - L\cos\alpha)^2]^2}\right) \\
& \left. + 16 b_1 b_2 L \sin\alpha \frac{(L\cos\alpha - h)(2h - L\cos\alpha)h}{[L^2\sin^2\alpha + (2h - L\cos\alpha)^2]^2}\right],
\end{aligned} \tag{5.21}$$

where $(b_1, b_2, b_3) = (b_1^p, b_2^p, b_3^p)$. The first two terms on the righthand side of (5.21) give the part of the interaction term (including the contribution from the mutual core boundary) that

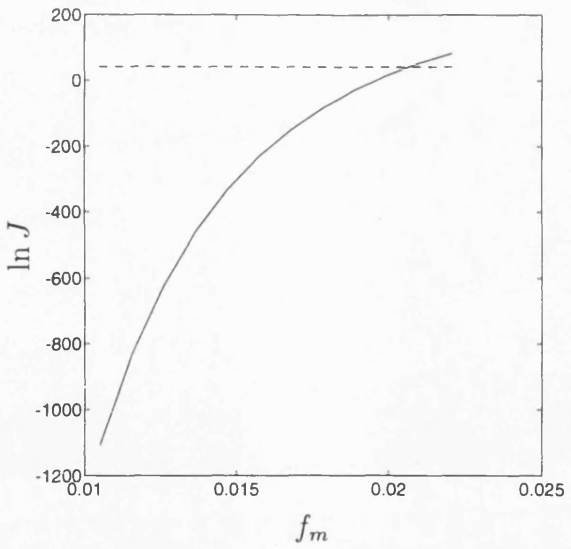

Figure 5.12: Natural logarithm of the nucleation rate for a complementary dislocation loop at the site of a pre-existing 60° dislocation, plotted as a function of fractional lattice mismatch f_m. The broken line marks the observability condition, $\ln J \approx 41$. The nucleation rate was put into units of $\mathrm{cc}^{-1}\mathrm{s}^{-1}$ before taking the logarithm.

arises from the infinite-body field of the pre-existing dislocation. This part has been documented by Gosling[24, equations (12) and (14)]. The collection of terms in the large square brackets gives the correction due to the presence of a free surface, and tends to zero as h becomes large. Although for reference the effects of the free surface at $x_2 = h$ have been included in writing (5.21), we present results only for the case of $h \to \infty$. The major features of the nucleation event considered are expected to be captured through incorporation of the near field of the dislocation. The predicted nucleation rate, obtained using (5.1), for the nucleation of a complementary 60° loop at the site of a pre-exisiting dislocation is plotted in Fig. 5.11 as a function of fractional lattice mismatch f_m. The temperature was taken to be 800K. Such a nucleation event converts the pre-existing dislocation into a 90° dislocation. Slight care is required when choosing the observability condition for this nucleation event. The nucleation event described here can only occur at the site of a pre-existing 60° dislocation. We assume (somewhat arbitrarily) that a nucleation event occuring within approximately 10Å of a pre-exisiting dislocation produces the configuration depicted in Fig. 5.10 before the loop reaches the critical size, and therefore such a nucleation event is governed by the rate just calculated. Conversely, we suppose that a nucleation event must occur within approximately 10Å of a pre-existing dislocation in oreder that it be governed by this activation energy. Therefore, for regions of the device within 10Å of a 60° dislocation, the rate is defined by (5.1) with the activation energy just calculated. A

rate of $10^{18}\text{cc}^{-1}\text{s}^{-1}$ corresponds to one nucleation event occuring every second within every $10\text{Å}\times10\text{Å}\times1\mu\text{m}$ volume. Therefore, if the rate governed by the activation energy for nucleation at a 60° dislocation exceeds this value, then a nucleation event should occur on each 1μm length of 60° dislocation every second, and hence most 60° dislocations will be converted to 90° dislocations almost immediately after they are formed at the interface. Once again the precise choice of observability condition does not strongly affect the conclusions obtained. The value $J = 10^{18}\text{cc}^{-1}\text{s}^{-1}$ is marked by the dashed line in Fig. 5.11. The observability condition is satisfied for a mismatch strain just larger than 0.02, which is in excellent agreement with the threshold at which dislocation distributions are found expermimentally to change from consisting predominantly of 60° dislocations to consisting predominantly of 90° dislocations. When $\text{Ge}_x\text{Si}_{1-x}$ layers are considered, the threshold mismatch is fractionally larger, due to the larger shear modulus ($\mu = 6.81 \times 10^{10}$Pa) in this material system, but is still very close to 0.02, in good agreement with experiment.

The value of f_m at which the mechanism for conversion of 60° to 90° dislocations switches on coincides very closely with the value of f_m at which the onset of homogeneous nucleation was found to occur in the previous section. Therefore, as the lattice mismatch passes through 0.02, there should be observed not only the changes to the dislocations distribution that are associated with the onset of homogeneous nucleation, which were described in the previous section, but also a change in dislocation type from 60° to 90°. Hence all the major changes in the dislocation distribution as f_m increases have been explained.

5.5 Closing Remarks

Before providing a general review of the conclusions of this chapter, it is worth making some remarks concerning non-linear core energies, and the choice of the pre-exponential factor appearing in (5.1).

The energy expression of (5.2), with E^c included, gives rigorously the energy in the system outside a cylindrical region of radius q surrounding the dislocation. Thus the energy within a region close to the core of the dislocation is not included. The core radius q has been taken in the above calculations to be equal to the magnitude, b, of Burgers vector. The reason for this choice is that it is expected that Hookes law will not be grossly violated at a distance of b or more from the dislocation[9], and that linear elasticity theory, leading to (5.2) may be used with some confidence. Within a distance q from the dislocation the strains (which are in fact singular at the dislocation core) become so large that a linear-elastic calculation of the energy

becomes meaningless. Some other technique must be used for this region, involving non-linear interatomic potentials or the like. The non-linear core energy is expected to be proportional to the length of dislocation line, and may thus be included through the introduction of a core energy parameter β in the numerator of the argument of the logarithm in (5.10) or (5.11). (This is equivalent to using a core radius $q = b/\beta$.) A similar term may be included in (2.44) or (2.55) if straight dislocations are considered. Two-dimensional calculations for straight dislocations have suggested values for β ranging from about 2.5 to 5[31, 53], which dictated our choice of $\beta = 4$ in Chapter 3. To the knowledge of the author, atomic calculations have not been carried out for three-dimensional loop configurations. Since it is far from clear what the relationship will be between the core energy of a long straight dislocation, and that of a small dislocation loop, we chose to neglect the non-linear core energy entirely in the above calculations. This approximation will tend to lead to an underestimate of the activation energies and hence to an overestimate of the nucleation rates. However, this error will be partially compensated for by the approximation that the dislocation is fully formed, which will tend to lead to an overestimate of the activation energies. In reality the dislocation will be dissociated, and its energy will thereby be reduced. The very good agreement that was obtained in the previous section between theoretical and experimental values for the threshold mismatch at which the dislocation distribution changes suggests that just such a cancellation of errors occurs. The introduction of a core parameter β equal to two or four destroys the agreement. A different approach to the problem, which would circumvent the uncertainties concerning the core energy and the extent of dissociation would be to move from the purely continuum-elastic picture of a dislocation to a Peierls-Nabarro model, in which the displacement discontinuity across the glide plane varies continuously over the glide plane, rather than undergoing a jump at the dislocation. The result of such an approach is that there is no longer a singularity at the core of the dislocation, and so the core cut-off procedure described above is made redundant. Other uncertainties are, of course, introduced, in particular concerning an appropriate description of the shear stress on the glide plane as a function of the relative displacement across the glide plane, and the complexity of the model is increased considerably. Such an approach to the nucleation of dislocation loops has recently been followed, in three dimensions, by Rice and Beltz[59].

Some discussion is merited of the expression, (5.1), for the nucleation rate J. The expression appearing in (5.1) for the nucleation rate has essentially been transported from studies of the rate of droplet nucleation in supersaturated vapours[42, 18]. Although the principles and equations involved in the study of droplet formation may be transferred in a very natural way to the nucleation of vacancy or interstitial aggregates, the correspondance with the nucleation

of glide loops is not so clear. It is well accepted that dislocations move via the propagation of double kinks, which nucleate on the dislocation and the separate by glide, causing motion of the dislocation[31]. (If a free surface is present then single kinks may also nucleate and play a role in the motion of the dislocation.) It is reasonable to suppose that, similarly, a loop expands and contracts via the nucleation of such double kinks: the loop radius changes size when a pair separates, glides around the loop, and self-annihilates. The correspondance between this mechanism of expansion and contraction and that on which droplet nucleation studies are based (a droplet changes size by the adsorption or evaporation of a single molecule at the droplet surface) merits investigation. Preliminary formulation of a model of loop nucleation based on kink motion promises interesting results. The central component of the model is to postulate the existence of a dislocation loop of some shape and then to postulate the formation, by thermal activation, of an incipient double kink, with the two members of the double kink initially separated by a, the kink-jump distance. At some generic stage, the separation, $s = na$, may either expand, to $(n+1)a$, or contract, to $(n-1)a$. An attempt frequency is selected and the process is modelled as a discrete random walk, with probability p for expansion, q for contraction, and $1-p-q$ for no change. Termination occurs either when the separation, s, contracts to zero (resulting in no change in loop radius) or when s expands to size Na, the current loop perimeter. This latter occurence results in an increase of the loop radius if the kink tends to expand the loop, and a decrease if the kink tends to contract the loop. If it is assumed that the nucleation of kinks occurs on a longer time scale than motion of a kink around the loop, then 'attempt frequency' for expansion or contraction of the loop is then the double kink nucleation frequency, and the kink random walk calculation provides probabilities for expansion, or shrinkage, or no change, for a loop of given radius. These probabilities then define a random walk through the space of loop radii. The kink calculation is not related to equilibrium, but it may be postulated that difference equations for the probability distribution of the *loop* size do admit a time-independent solution, also obtainable from standard equilibrium theory via a Boltzmann term involving the loop energy as a function of radius (taken to be the elastic energy). If performed *just* at the level of the loops, this postulate provides the *ratio* of the probabilities for expansion and contraction of the loop radius as a function of known elastic energies. If p and q are assumed constant for a given loop size (an assumption that can be rationalized under certain circumstances), then these expansion and contraction probabilities for the loop radius can be independently related to the jump probabilities for the kinks, and are in fact just functions of the ratio p/q under the approximation described earlier regarding the relative timescales of kink nucleation and propagation. Therefore, the requirement that an appropriate steady-state solution exist, for the distribution

of loop radii, yields an equation that defines the ratio p/q, and hence defines separately the expansion and contraction probabilities for the loop. Additional information thus results from considering the kink motion. Furthermore, kink motion involves very few atoms and so attempt frequency for such motion is an easier concept to develop than the corresponding one for a loop, where growth involves the cooperation of many atoms, the exact number depending on the loop size. Preliminary results of an approximate development of the model suggest that the nucleation rate is governed by an expression of the type of (5.1) but with a much lower pre-exponential factor, of the order of 10^{20} as opposed to 10^{36}. The threshold values of lattice mismatch at which homogeneous nucleation and 90° dislocation formation are predicted to occur are defined almost entirely by the exponential dependence of the nucleation rate on the activation energy. Even changing the pre-exponential factor to 10^{20} yields a value of 0.022 for threshold lattice mismatch at which the onset of homogeneous nucleation should occur. This value is very close to the value of 0.021 obtained using a pre-exponential factor of 10^{36}. However, the predicted rates of nucleation once the threshold lattice mismatch is exceeded are much more sensitive to the choice of pre-exponential factor, as is illustrated in Fig. 5.10. When the threshold lattice mismatch is exceeded, the activation energy for nucleation quite rapidly approaches zero, which corresponds to spontaneous emmision of loops at a rate given by the pre-exponential factor. The areal nucleation rate obtained when the pre-exponential is 10^{36} yields predicted threading segment densities that are very many orders of magnitude greater than those that are observed, as illustrated by the solid curve in Fig. 5.10. However, when the smaller pre-exponential factor of 10^{20} is used then the predicted threading segment densities are in excellent agreement with experiment, as illustrated by the dashed curve in Fig. 5.10. This finding encourages further investigation of the loop nucleation model described above.

We now summarize briefly the findings of this chapter. Using the representation for the stresses due to an arbitrary dislocation in a half-space that was developed in Chapter 4, an estimate of the threshold lattice mismatch at which homogeneous nucleation at the free surface becomes possible has been obtained. This threshold value was found to be 0.021. Homogeneous nucleation of loops buried within the layer was found to be negligible for the range of mismatch considered ($f_m < 0.035$). It was found that use of approximations based on infinite-body results leads to large errors: in particular, these approximations yield a predicted value of 0.037 for the threshold mismatch marking the onset of homogeneous nucleation at the free surface, as opposed to the value of 0.021 obtained using the exact results. By considering the implications of a shift from heterogeneous to predominantly homogeneous nucleation, the change in dislocation distribution that is observed to take place over the mismatch range 0.015

to 0.025 was explained. Quantitative study of the increase in threading segment density over this mismatch range suggested that a pre-exponential of around 10^{20} in the rate expression, (5.1), is appropriate, as is supported by preliminary investigations into a model of loop nucleation based on kink motion. The change in dislocation *type* that is observed to occur over the same mismatch range was explained by considering a nucleation event that converts a 60° dislocation into a 90° dislocation, and by illustrating that this event should occur on an experimentally significant scale at a mismatch strain of around 0.02. In this way a fairly thorough qualitative and partially quantitative description of dislocation distributions, which exhibits good agreement with experiment, has been provided for $f_m < 0.035$.

Chapter 6

The Stresses due to a Periodic Array of Two-Dimensional Inclusions

The stresses due to the lattice mismatch in a layer that has a planar interface with its substrate are given in Section 2.2. However, devices are often constructed on a patterned substrate[62, 21, 19] (one that has been etched prior to growth of the layer) or on an off-cut substrate[1, 73, 47]. In such cases the interface between the layer and substrate is no longer planar, and non-uniformities in the stresses may arise. The non-uniformites may be important both because of the effects they have on the electronic properties of the device and because of the effect they have on its mechanical stability: localized regions of high stress facilitate the nucleation of defects such as dislocations, which can degrade device performance. When considering aspects of the nucleation process, a solution for the stresses due to a non-planar interface, which we derive in this chapter, will be most useful. As will be shown shortly, the basic solution that is required for determining the stresses due to a non-planar interface is that for the stresses due to a periodic array of two-dimensional mismatched inclusions. It is this solution that we obtain in Sections 6.1 through 6.3. The solution is yet to be fully exploited, but in Section 6.4 we provide illustrations of its utility by calculating activation energies for loop formation at typical non-planar interface features. As a separate application, it is noted that the stresses due to an array of strained buried quantum wires may be determined by modelling the wires as inclusions. In Section 6.5 we use the solution obtained to assess the strain-induced band gap shift in strained quantum

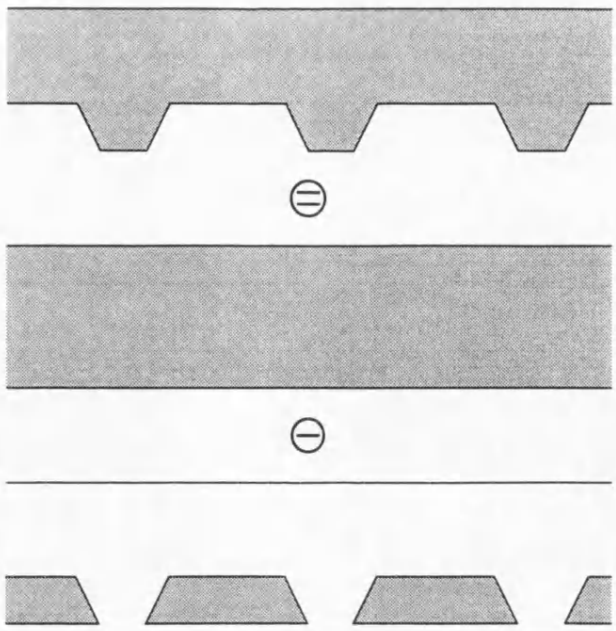

Figure 6.1: Schematic illustration of a typical interface between a strained layer and patterned substrate (top). Also illustrated is the decomposition, whereby the configuration of interest may be constructed as the superposition of a strained layer with a planar interface and a periodic array of inclusions. The $\ominus$ sign means that the stresses due to the inclusions must be subtracted from those due to the layer with a planar interface.

wire arrays. The results obtained show very favourable agreement with experiment.

A typical substrate-layer interface that is non-planar is illustrated in Fig. 6.1. The shaded regions are the layer material, whose lattice is mismatched with that of the substrate. Also shown in this figure is the manner in which the configuration of interest may be represented as the superposition of a strained layer with a planar interface, and a periodic array of two-dimensional inclusions. The resultant stresses may therefore be written as the sum of the stresses of the component configurations. (The $\ominus$ sign means that the stresses due to the periodic array of inclusions must be subtracted from those due to the strained layer with a planar interface. This is because the array of inclusions is, in effect, used to 'remove' areas of mismatch from the first component configuration.) The stresses in a layer with a planar interface are given in Section 2.2. All that remains is to assess stresses due to the periodic array of inclusions.

6.1 Derivation of the Stresses

The geometry of the most general system that we shall require and the coordinate system employed are depicted in Fig. 6.2. We consider a periodic array of quadrilateral, two-

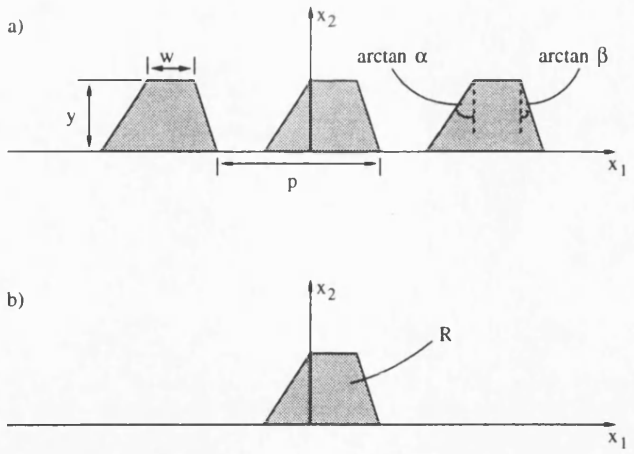

Figure 6.2: (a) Schematic illustration of the periodic array of inclusions, defining angles, dimensions, and coordinates. (b) The case $p \to \infty$: a single inclusion R.

dimensional inclusions. The system is unbounded in the x_1- and x_3-directions and is periodic with period p in the direction of x_1, but there is no variation with respect to x_3; the system is therefore in a state of generalized plane strain. The quantities α, β, w, and y are defined in Fig. 6.2a. The inclusions are of a material whose lattice fails to match that of the surrounding material by an amount defined through the mismatch-strain tensor e^m_{ij}. This is defined so that if an inclusion is subjected to a strain $-e^m_{ij}$, the lattice in that inclusion is then brought into conformity with that of the surrounding material. The strains are assumed small enough for the linear theory of elasticity to apply. With displacement $\mathbf{u}$ measured relative to the configuration in which all mismatched regions are subjected to a strain $-e^m_{ij}$, the stress-strain relation for the body is therefore

$$\sigma_{ij} = c_{ijkl}(e_{kl} - e^m_{kl}), \tag{6.1}$$

where c_{ijkl}, σ_{ij}, and e_{ij} represent, respectively, components of the tensors of elastic moduli, stress, and strain. In order that (6.1) apply throughout the body, the tensor e^m_{ij} is defined to be zero outside the mismatched regions. We shall only consider a bimaterial, so e^m_{ij} will be constant within the inclusions, and hence piecewise-constant throughout the body; specifically,

$$e^m_{ij} = \epsilon^m_{ij}\chi_\Omega, \tag{6.2}$$

where ϵ_{ij}^m is a constant tensor, and χ_Ω is the characteristic function of the mismatched regions, denoted Ω. Moreover, only the case of elastic isotropy will be considered, so that

$$c_{ijkl} = \frac{2\mu\nu}{1-2\nu}\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}).$$

Here μ is the shear modulus of the material, and ν is Poisson's ratio, each assumed constant throughout the body. In most applications devices are grown on a thick substrate and it is justifiable to model the region $x_2 < 0$ as a half-space. (In particular, no mean strain in the substrate will be allowed, so displacements will be constrained to remain bounded at the lateral edges of the sample.) However, in general there will be a free surface at $x_2 = h$, for some h. If this surface is *flat* then it is in principle quite straightforward to allow for its effects, although there results a considerable increase in the complexity of the formulae obtained. If the surface is non-planar, then it is not possible to account exactly for its effects without computation. Our approach will be to allow for the major effect of the free-surface, by allowing the stresses σ_{i2} to relax in the mean, whilst disregarding the detail of the image stresses. Application of the formulae will, correspondingly, be restricted to cases in which such detail can be expected to be insignificant; such an expectation is fulfilled provided h is sufficiently large.

6.2 Evaluation of the Stresses

It is convenient first to consider the case of a single inclusion, R: i.e. the case $p \to \infty$, illustrated in Fig. 6.2b. Because the system is in a state of generalized plane strain, the equilibrium equations become

$$c_{i\alpha k\beta}u_{k,\alpha\beta} - c_{i\alpha kl}e^m_{kl,\alpha} = 0. \tag{6.3}$$

The conditions to be imposed are that the stresses tend to zero at infinity[1] and that tractions are continuous across any internal surface of the body. The convention has been adopted that while repeated Roman subscripts sum over the values 1, 2, 3, repeated Greek subscripts sum just over the values 1, 2. The term $e^m_{kl,\alpha}$, being the derivative of a piecewise-constant function, represents a generalized function. Given a Green's tensor obeying

$$c_{i\alpha k\beta}G_{mk,\alpha\beta}(\mathbf{x}) + \delta_{im}\delta(\mathbf{x}) = 0, \tag{6.4}$$

[1] If the body is finite then the tractions are required to be zero on the boundary of the body. However, since we are neglecting the detailed effects of the boundary of the body, it is appropriate to apply infinite-body boundary conditions and, in the sequel, to use the infinite-body Green's tensor.

where $\mathbf{x} = (x_1, x_2)$, the solution to (6.3) may be written

$$\begin{aligned} u_k^s(\mathbf{x}) &= -\int dx' G_{mk}(\mathbf{x}-\mathbf{x}') c_{m\alpha rl} e^m_{rl,\alpha}(\mathbf{x}') \\ &= \int dx' \frac{\partial}{\partial x'_\alpha} G_{mk}(\mathbf{x}-\mathbf{x}') c_{m\alpha rl} e^m_{rl}(\mathbf{x}'). \end{aligned}$$

In the above dx' represents $dx'_1 dx'_2$ and integrations take place over the whole plane. The superscript 's' has been used to emphasize that a single inclusion is being considered. The second equality has been obtained by integrating by parts and observing that e^m_{ij} has bounded support. If, for example, the effects of a flat free surface were to be incorporated rigorously, then G_{mk} would be replaced by the half-space Green's tensor in the above expression. Since, for an infinite body, G_{mk} is a function of $\mathbf{x} - \mathbf{x}'$, and e^m_{ij} is ϵ^m_{ij} in R and zero otherwise, the displacement may be written

$$u_k^s(\mathbf{x}) = -\int_R dx' G_{mk,\alpha}(\mathbf{x}-\mathbf{x}') c_{m\alpha rl} \epsilon^m_{rl},$$

which is essentially Eshelby's standard solution[16]. The stresses are obtained trivially from (6.1):

$$\sigma^s_{pq}(\mathbf{x}) = -c_{pqr\beta} c_{i\alpha kl} \epsilon^m_{kl} \int_R dx' G_{ir,\alpha\beta}(\mathbf{x}-\mathbf{x}') - c_{pqkl} \epsilon^m_{kl} \chi_R(\mathbf{x}). \tag{6.5}$$

Because the problem is linear, the stresses due to the periodic array of inclusions may be obtained by summing appropriately the stresses due to a single inclusion. Using an 'a' superscript to mark the stresses due to a periodic array of inclusions we obtain

$$\sigma^a_{pq}(\mathbf{x}) = \sum_{n=-\infty}^{\infty} \sigma^s_{pq}(x_1 - np, x_2),$$

or, using a Fourier-series representation,

$$\sigma^a_{pq}(\mathbf{x}) = \frac{1}{p} \sum_{\substack{k=-\infty \\ \xi = 2\pi k/p}}^{\infty} \tilde{\sigma}^s_{pq}(\xi, x_2) e^{-i\xi x_1}, \tag{6.6}$$

where $\tilde{\sigma}^s$ represents the Fourier transform of σ^s:

$$\tilde{\sigma}^s(\xi, x_2) = \int dx_1 \sigma^s(x_1, x_2) e^{i\xi x_1}.$$

Note that, since σ^s must be real, $\tilde{\sigma}^s(-\xi, x_2) = \overline{\tilde{\sigma}^s(\xi, x_2)}$, and so (6.6) may be written

$$\sigma^a_{pq}(\mathbf{x}) = \frac{2}{p}\mathrm{Re} \sum_{\substack{k=1 \\ \xi=2\pi k/p}}^{\infty} \tilde{\sigma}^s_{pq}(\xi, x_2) e^{-\mathrm{i}\xi x_1} + \bar{\sigma}^a_{pq}(x_2), \tag{6.7}$$

where

$$\bar{\sigma}^a_{pq}(x_2) = \frac{1}{p}\int_0^p dx_1 \sigma^a_{pq}(x_1, x_2)$$

are the mean values, with respect to x_1, of the stresses. A short overbar on any quantity will represent the mean of that quantity, as defined above for σ^a_{pq}. Inserting expression (6.5) for σ^s into (6.7), performing and inverting a Fourier transform with respect to x_2, and interchanging the order of integration where necessary, we obtain

$$\begin{aligned} \sigma^a_{pq}(\mathbf{x}) \;=\; & c_{pqr\beta} c_{i\alpha kl} \epsilon^m_{kl} \frac{1}{\pi p}\mathrm{Re} \sum_{\substack{k=1 \\ \xi_1=2\pi k/p}}^{\infty} \left(e^{-\mathrm{i}\xi_1 x_1} \int d\xi_2 e^{-\mathrm{i}\xi_2 x_2} \xi_\alpha \xi_\beta \hat{G}_{ir}(\boldsymbol{\xi}) \int_R dx' e^{\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}'} \right) \\ & - c_{pqkl}\epsilon^m_{kl}(\chi_{\cup R}(\mathbf{x}) - \bar{\chi}_{\cup R}(x_2)) + \bar{\sigma}^a_{pq}(x_2). \end{aligned} \tag{6.8}$$

Here $\chi_{\cup R}(\mathbf{x})$ represents the characteristic function of the periodic array of mismatched regions, denoted $\cup R$, and $\hat{G}(\boldsymbol{\xi})$ is the two-dimensional Fourier transform of $G(\mathbf{x})$:

$$\hat{G}(\boldsymbol{\xi}) = \int\int dx_1 dx_2 G(\mathbf{x}) e^{\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}}.$$

The integral over the region R that appears in (6.8) is trivial, and the integral with respect to ξ_2 may be evaluated using standard complex variable methods. The series may then be summed explicitly. Finally, the mean values of the stresses may be determined in a manner very similar to that employed in Section 2.2. To see this, first note that from (6.1),

$$\bar{\sigma}^a_{ij} = c_{ijkl}(\bar{e}^a_{kl} - \bar{e}^m_{kl}),$$

where $e^m_{kl}(\mathbf{x})$ is given by (6.2) with $\Omega = \cup R$, and the short overbar represents the mean value, over a period p, with respect to x_1. Now

$$\bar{e}^a_{kl} = \frac{1}{2p}\int_0^p dx_1 (u^a_{k,l} + u^a_{l,k}).$$

The displacement $\mathbf{u}^a$ is clearly continuous, and is periodic in x_1, because no mean deformation can be induced in the infinite substrate; moreover, $\frac{\partial}{\partial x_3} \equiv 0$ because the problem is two-

dimensional. Therefore,

$$\bar{e}^a_{13} = \bar{e}^a_{11} = \bar{e}^a_{33} = 0.$$

Since only $\bar{e}^a_{i2}$ are non-zero, the reasoning now follows exactly that of Section 2.2 with, as before, the stresses $\bar{\sigma}^a_{i2}$ being taken equal to zero identically. Hence,

$$\bar{\sigma}^a_{pq} = -\Gamma_{pqkl}\bar{e}^m_{kl},$$

where Γ_{ijkl} is given in (3.7). For isotropy this reduces to the following non-zero stresses:

$$\bar{\sigma}^a_{11} = \frac{-2\mu}{1-\nu}(\bar{e}^m_{11} + \nu\bar{e}^m_{33}), \quad \bar{\sigma}^a_{33} = \frac{-2\mu}{1-\nu}(\nu\bar{e}^m_{11} + \bar{e}^m_{33}), \quad \bar{\sigma}^a_{13} = -2\mu\bar{e}^m_{13}. \tag{6.9}$$

From (6.2) it is immediate that

$$\bar{e}^m_{kl}(x_2) = \begin{cases} \frac{(\alpha+\beta)(y-x_2)+w}{p}\epsilon^m_{kl} & \text{for } 0 < x_2 < y \\ 0 & \text{otherwise.} \end{cases}$$

The only term that remains to be evaluated in (6.8) is henceforth denoted τ_{pq}:

$$\tau_{pq}(\mathbf{x}) = c_{pqr\beta}c_{i\alpha kl}\epsilon^m_{kl}\frac{1}{\pi p}\mathrm{Re}\sum_{\substack{k=1 \\ \xi_1=2\pi k/p}}^{\infty}\left(e^{-\mathrm{i}\xi_1 x_1}\int d\xi_2 e^{-\mathrm{i}\xi_2 x_2}\xi_\alpha\xi_\beta\hat{G}_{ir}(\boldsymbol{\xi})\int_R dx' e^{\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}'}\right). \tag{6.10}$$

The integral over the region R may be performed directly; with reference to the geometry depicted in Fig. 6.2 we obtain

$$\int_R dx' e^{\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}'} = \frac{e^{\mathrm{i}\xi_1(w_1+\beta y)}}{\xi_1(\xi_2-\xi_1\beta)}\left(1-e^{\mathrm{i}(\xi_2-\xi_1\beta)y}\right) - \frac{e^{-\mathrm{i}\xi_1\alpha y}}{\xi_1(\xi_2+\xi_1\alpha)}\left(1-e^{\mathrm{i}(\xi_2+\xi_1\alpha)y}\right).$$

Now define

$$F_{\alpha\beta ij}(z,\gamma,\xi_1) = \int d\xi_2 \frac{\xi_\alpha\xi_\beta\hat{G}_{ij}(\boldsymbol{\xi})e^{-\mathrm{i}\xi_2 z}}{\xi_2+\gamma\xi_1}. \tag{6.11}$$

Then (6.10) may be written

$$\begin{aligned}\tau_{pq}(\mathbf{x}) \;=\; & c_{pqj\beta}c_{i\alpha kl}\epsilon^m_{kl}\frac{1}{\pi p}\mathrm{Re}\sum_{\substack{k=1 \\ \xi_1=2\pi k/p}}^{\infty}\frac{e^{-\mathrm{i}\xi_1 x_1}}{\xi_1} \\ & \times\Big[F_{\alpha\beta ij}(x_2,-\beta,\xi_1)e^{\mathrm{i}\xi_1(w+\beta y)} - F_{\alpha\beta ij}(x_2-y,-\beta,\xi_1)e^{\mathrm{i}\xi_1 w} \\ & + F_{\alpha\beta ij}(x_2-y,\alpha,\xi_1) - F_{\alpha\beta ij}(x_2,\alpha,\xi_1)e^{-\mathrm{i}\xi_1\alpha y}\Big].\end{aligned} \tag{6.12}$$

The two-dimensional Fourier transform of the plane Green's tensor, satisfying (6.4), is precisely the three-dimensional transform of the usual whole-space Green's tensor with $\xi_3 \equiv 0$, because of the generalized plane-strain condition. The required transform is therefore[14]

$$\hat{G}_{ij}(\boldsymbol{\xi}) = \frac{1}{\mu}\left(\frac{\delta_{ij}}{|\boldsymbol{\xi}|^2} - \frac{1}{2(1-\nu)}\frac{\xi_i\xi_j}{|\boldsymbol{\xi}|^4}\right), \quad \xi_3 \equiv 0. \tag{6.13}$$

If we define

$$I(z,\gamma,\xi_1) = \frac{1}{\mu}\int d\xi_2 \frac{1}{|\boldsymbol{\xi}|^2}\frac{e^{-\mathrm{i}\xi_2 z}}{\xi_2+\gamma\xi_1},$$

and

$$J(z,\gamma,\xi_1) = \frac{1}{2\mu(1-\nu)}\int d\xi_2 \frac{1}{|\boldsymbol{\xi}|^4}\frac{e^{-\mathrm{i}\xi_2 z}}{\xi_2+\gamma\xi_1},$$

where the integrals are to be interpreted as Cauchy principal values, then it is immediate from (6.11) and (6.13) that

$$F_{\alpha\beta ij}(z,\gamma,\xi_1) = \delta_{ij}\xi_1^{(2-m)}\left(\mathrm{i}\frac{\partial}{\partial z}\right)^m I(z,\gamma,\xi_1) - (1-\delta_{i3})(1-\delta_{j3})\xi_1^{(4-n)}\left(\mathrm{i}\frac{\partial}{\partial z}\right)^n J(z,\gamma,\xi_1). \tag{6.14}$$

Here

$$m = \delta_{\alpha 2} + \delta_{\beta 2} \text{ and } n = \delta_{\alpha 2} + \delta_{\beta 2} + \delta_{i2} + \delta_{j2}. \tag{6.15}$$

The integrals I and J are easily evaluated using complex-variable methods, as is described in Appendix E. For $z > 0$ and $\xi_1 > 0$ the results are

$$I(z,\gamma,\xi_1) = \frac{\pi}{\mu\xi_1^2}\left(\frac{e^{-\xi_1 z}}{\gamma-\mathrm{i}} - \frac{\mathrm{i}e^{\mathrm{i}\xi_1\gamma z}}{\gamma^2+1}\right),$$

and

$$J(z,\gamma,\xi_1) = \frac{\pi}{4\mu(1-\nu)\xi_1^4}\left[\left(\frac{\gamma-2\mathrm{i}}{(\gamma-\mathrm{i})^2} + \frac{\xi_1 z}{\gamma-\mathrm{i}}\right)e^{-\xi_1 z} - \frac{2\mathrm{i}e^{\mathrm{i}\xi_1\gamma z}}{(\gamma^2+1)^2}\right].$$

Differentiation as required by (6.14) yields the following explicit formula for $F_{\alpha\beta ij}$ when $z > 0$ and $\xi_1 > 0$:

$$F_{\alpha\beta ij}(z,\gamma,\xi_1) = \delta_{ij}\frac{\pi}{\mu}\left((-\mathrm{i})^m\frac{e^{-\xi_1 z}}{\gamma-\mathrm{i}} - \mathrm{i}(-\gamma)^m\frac{e^{\mathrm{i}\xi_1\gamma z}}{\gamma^2-1}\right)$$

$$+\frac{(1-\delta_{i3})(1-\delta_{j3})\pi}{4\mu(1-\nu)}\left[\left(\frac{n}{\gamma-\mathrm{i}} - \frac{\gamma-2\mathrm{i}}{(\gamma-\mathrm{i})^2} - \frac{\xi_1 z}{\gamma-\mathrm{i}}\right)(-\mathrm{i})^n e^{-\xi_1 z} + \frac{2\mathrm{i}(-\gamma)^n}{(\gamma^2+1)^2}e^{\mathrm{i}\xi_1\gamma z}\right],$$

where m and n are defined in (6.15). From (6.11) it is clear that

$$F_{\alpha\beta ij}(-z,\gamma,\xi_1) = \overline{F_{\alpha\beta ij}(z,\gamma,\xi_1)}$$

and thus the case of $z < 0$ is dealt with simply. To deal with the summation required by (6.12), we first collect like terms in $F_{\alpha\beta ij}$ by writing, for $z > 0$,

$$F_{\alpha\beta ij}(z,\gamma,\xi_1) = A_{\alpha\beta ij}(\gamma)e^{-\xi_1 z} + B_{\alpha\beta ij}(\gamma)z\xi_1 e^{-\xi_1 z} + C_{\alpha\beta ij}(\gamma)e^{\mathrm{i}\xi_1\gamma z},$$

where

$$\begin{aligned}
A_{\alpha\beta ij}(\gamma) &= \delta_{ij}\frac{(-\mathrm{i})^m\pi}{\mu(\gamma-\mathrm{i})} - (1-\delta_{i3})(1-\delta_{j3})\frac{(-\mathrm{i})^n\pi}{4\mu(1-\nu)}\left(\frac{\gamma-2\mathrm{i}}{(\gamma-\mathrm{i})^2} - \frac{n}{\gamma-\mathrm{i}}\right), \\
B_{\alpha\beta ij}(\gamma) &= -(1-\delta_{i3})(1-\delta_{j3})\frac{(-\mathrm{i})^n\pi}{4\mu(1-\nu)}\frac{1}{\gamma-\mathrm{i}}, \\
C_{\alpha\beta ij}(\gamma) &= -\mathrm{i}\delta_{ij}(-\gamma)^m\frac{\pi}{\mu}\frac{1}{\gamma^2+1} + (1-\delta_{i3})(1-\delta_{j3})\frac{\mathrm{i}(-\gamma)^n\pi}{2\mu(1-\nu)}\frac{1}{(\gamma^2+1)^2}.
\end{aligned} \tag{6.16}$$

Next, writing

$$\begin{aligned}
K_{pq} &= c_{pqj\beta}c_{i\alpha kl}\epsilon^m_{kl}A_{\alpha\beta ij}, \\
L_{pq} &= c_{pqj\beta}c_{i\alpha kl}\epsilon^m_{kl}B_{\alpha\beta ij}, \\
M_{pq} &= c_{pqj\beta}c_{i\alpha kl}\epsilon^m_{kl}C_{\alpha\beta ij},
\end{aligned} \tag{6.17}$$

it is straightforward to show that (6.12) takes the form

$$\tau_{pq}(\mathbf{x}) = \frac{1}{\pi p}\mathrm{Re}\sum_{\substack{k=1\\ \xi=2\pi k/p}}^{\infty} T_{pq}(\xi), \tag{6.18}$$

where for $x_2 > y$

$$\begin{aligned}
T_{pq}(\xi) &= K_{pq}(-\beta)\frac{e^{-\xi[|x_2|+\mathrm{i}(x_1-w-\beta y)]}}{\xi} + L_{pq}(-\beta)\,|x_2|\,e^{-\xi[|x_2|+\mathrm{i}(x_1-w-\beta y)]} \\
&\quad -K_{pq}(-\beta)\frac{e^{-\xi[|x_2-y|+\mathrm{i}(x_1-w)]}}{\xi} - L_{pq}(-\beta)\,|x_2-y|\,e^{-\xi[|x_2-y|+\mathrm{i}(x_1-w)]} \\
&\quad +K_{pq}(\alpha)\frac{e^{-\xi(|x_2-y|+\mathrm{i}x_1)}}{\xi} + L_{pq}(\alpha)\,|x_2-y|\,e^{-\xi(|x_2-y|+\mathrm{i}x_1)} \\
&\quad -K_{pq}(\alpha)\frac{e^{-\xi[|x_2|+\mathrm{i}(x_1+\alpha y)]}}{\xi} - L_{pq}(\alpha)\,|x_2|\,e^{-\xi[|x_2|+\mathrm{i}(x_1+\alpha y)]},
\end{aligned}$$

which also holds for $x_2 < 0$ if K_{pq} and L_{pq} are replaced by their complex conjugates, and for

$0 < x_2 < y$

$$
\begin{aligned}
T_{pq}(\xi) \;=\; & K_{pq}(-\beta)\frac{e^{-\xi[x_2+\mathrm{i}(x_1-w-\beta y)]}}{\xi} + L_{pq}(-\beta)x_2 e^{-\xi[x_2+\mathrm{i}(x_1-w-\beta y)]} \\
& -\overline{K_{pq}(-\beta)}\frac{e^{-\xi[|x_2-y|+\mathrm{i}(x_1-w)]}}{\xi} - \overline{L_{pq}(-\beta)}\,|x_2-y|\,e^{-\xi[|x_2-y|+\mathrm{i}(x_1-w)]} \\
& + \left(M_{pq}(-\beta) - \overline{M_{pq}(-\beta)}\right)\frac{e^{\mathrm{i}\xi[-\beta(x_2-y)-x_1+w]}}{\xi} \\
& +\overline{K_{pq}(\alpha)}\frac{e^{-\xi(|x_2-y|+\mathrm{i}x_1)}}{\xi} + \overline{L_{pq}(\alpha)}\,|x_2-y|\,e^{-\xi(|x_2-y|+\mathrm{i}x_1)} \\
& -K_{pq}(\alpha)\frac{e^{-\xi[x_2+\mathrm{i}(x_1+\alpha y)]}}{\xi} - L_{pq}(\alpha)x_2 e^{-\xi[x_2+\mathrm{i}(x_1+\alpha y)]} \\
& - \left(M_{pq}(\alpha) - \overline{M_{pq}(\alpha)}\right)\frac{e^{\mathrm{i}\xi[\alpha(x_2-y)-x_1]}}{\xi}.
\end{aligned}
$$

All that remains is to sum the series, as required by (6.18). The summation of terms multiplied by the coefficients K_{pq} and L_{pq} involves no difficulty: each involves a term that decays exponentially, and knowledge of the following simple series is sufficient:

$$
\sum_{\substack{k=1 \\ \xi=2\pi k/p}}^{\infty} e^{-\xi\phi} = \frac{e^{2\pi\phi/p}}{1-e^{2\pi\phi/p}}
$$

gives the sum of a geometric series, and integration with respect to ϕ yields

$$
\sum_{\substack{k=1 \\ \xi=2\pi k/p}}^{\infty} \frac{e^{-\xi\phi}}{\xi} = \frac{p}{2\pi}\ln\left(\frac{1}{1-e^{-2\pi\phi/p}}\right),
$$

where $\mathrm{Re}(\phi) > 0$. Now, from (6.16) and (6.17) it is clear that M_{pq} is purely imaginary; therefore the other series required take the form

$$
\begin{aligned}
\frac{2}{p}\mathrm{Re}\sum_{\substack{k=1 \\ \xi=2\pi k/p}}^{\infty} M_{pq}\frac{e^{\mathrm{i}\xi\phi}}{\xi} &= \frac{1}{p}\sum_{\substack{k=1 \\ \xi=2\pi k/p}}^{\infty}\left(M_{pq}\frac{e^{\mathrm{i}\xi\phi}}{\xi} + \overline{M_{pq}}\frac{e^{-\mathrm{i}\xi\phi}}{\xi}\right) \\
&= \frac{1}{p}\sum_{\substack{k=1 \\ \xi=2\pi k/p}}^{\infty} M_{pq}\left(\frac{e^{\mathrm{i}\xi\phi}}{\xi} + \frac{e^{-\mathrm{i}\xi\phi}}{-\xi}\right) \\
&= \frac{\mathrm{i}}{p}M_{pq}\sum_{\substack{k=-\infty \\ k\neq 0,\ \xi=2\pi k/p}}^{\infty}\frac{\mathrm{i}}{\xi}e^{-\mathrm{i}\xi\phi}. \qquad (6.19)
\end{aligned}
$$

Now, consider the function

$$f(\phi) = \frac{1}{2}\mathrm{sgn}\phi - \frac{\phi}{p}, \quad -p/2 < \phi < p/2.$$

The Fourier coefficient of f is

$$\int_{-p/2}^{p/2} f(\phi) e^{\mathrm{i}\xi\phi},$$

which is found to be

$$\frac{\mathrm{i}}{\xi}\left(1 - \frac{2\sin(\xi p/2)}{\xi p}\right);$$

this is 0 when $\xi = 0$ and is equal to i/ξ when $\xi = 2\pi k/p$, where k is any integer different from zero. Therefore the infinite sum appearing on the right-hand side of the last equality in (6.19) is simply the p-periodic extension of the function f; hence (6.19) gives

$$\frac{2}{p}\mathrm{Re}\sum_{\substack{k=1 \\ \xi=2\pi k/p}}^{\infty} M_{pq}\frac{e^{\mathrm{i}\xi\phi}}{\xi} = \mathrm{i}M_{pq}\left(\frac{1}{2}\sum_{n=-\infty}^{\infty}\mathrm{sgn}(\phi - np) - \frac{\phi}{p}\right),$$

where

$$\sum_{n=-\infty}^{\infty}\mathrm{sgn}(\phi - np) = \lim_{N\to\infty}\sum_{n=-N}^{N}\mathrm{sgn}(\phi - np).$$

This is the final series required. Summation of the series in (6.18) yields: for $x_2 > y$

$$\begin{aligned}
\tau_{pq}(\mathbf{x}) \quad = \frac{1}{\pi p}\mathrm{Re} \quad \Bigg\{ & - K_{pq}(-\beta)\frac{p}{2\pi}\ln\left(1 - e^{-2\pi[|x_2| + \mathrm{i}(x_1 - w - \beta y)]/p}\right) \\
& + L_{pq}(-\beta)\,|x_2|\,\frac{e^{-2\pi[|x_2| + \mathrm{i}(x_1 - w - \beta y)]/p}}{1 - e^{-2\pi[|x_2| + \mathrm{i}(x_1 - w - \beta y)]/p}} \\
& + K_{pq}(-\beta)\frac{p}{2\pi}\ln\left(1 - e^{-2\pi[|x_2 - y| + \mathrm{i}(x_1 - w)]/p}\right) \\
& - L_{pq}(-\beta)\,|x_2 - y|\,\frac{e^{-2\pi[|x_2 - y| + \mathrm{i}(x_1 - w)]/p}}{1 - e^{-2\pi[|x_2 - y| + \mathrm{i}(x_1 - w)]/p}} \\
& - K_{pq}(\alpha)\ln\left(1 - e^{-2\pi[|x_2 - y| + \mathrm{i}x_1]/p}\right) \\
& + L_{pq}(\alpha)\,|x_2 - y|\,\frac{e^{-2\pi[|x_2 - y| + \mathrm{i}x_1]/p}}{1 - e^{-2\pi[|x_2 - y| + \mathrm{i}x_1]/p}} \\
& + K_{pq}(\alpha)\frac{p}{2\pi}\ln\left(1 - e^{-2\pi[|x_2| + \mathrm{i}(x_1 + \alpha y)]/p}\right) \\
& - L_{pq}(\alpha)\,|x_2|\,\frac{e^{-2\pi[|x_2| + \mathrm{i}(x_1 + \alpha y)]/p}}{1 - e^{-2\pi[|x_2| + \mathrm{i}(x_1 + \alpha y)]/p}}\Bigg\},
\end{aligned} \tag{6.20}$$

which also holds for $x_2 < 0$ if K_{pq} and L_{pq} are replaced by their complex conjugates, and for

$0 < x_2 < y$

$$
\begin{aligned}
\tau_{pq}(\mathbf{x}) \quad = \frac{1}{\pi p}\mathrm{Re} \quad & \Bigg\{ - K_{pq}(-\beta)\frac{p}{2\pi}\ln\left(1 - e^{-2\pi[|x_2|+\mathrm{i}(x_1-w-\beta y)]/p}\right) \\
& + L_{pq}(-\beta)\,|x_2|\,\frac{e^{-2\pi[|x_2|+\mathrm{i}(x_1-w-\beta y)]/p}}{1 - e^{-2\pi[|x_2|+\mathrm{i}(x_1-w-\beta y)]/p}} \\
& + \overline{K_{pq}(-\beta)}\frac{p}{2\pi}\ln\left(1 - e^{-2\pi[|x_2-y|+\mathrm{i}(x_1-w)]/p}\right) \\
& - \overline{L_{pq}(-\beta)}\,|x_2-y|\,\frac{e^{-2\pi[|x_2-y|+\mathrm{i}(x_1-w)]/p}}{1 - e^{-2\pi[|x_2-y|+\mathrm{i}(x_1-w)]/p}} \\
& - \overline{K_{pq}(\alpha)}\ln\left(1 - e^{-2\pi[|x_2-y|+\mathrm{i}x_1]/p}\right) \\
& + \overline{L_{pq}(\alpha)}\,|x_2-y|\,\frac{e^{-2\pi[|x_2-y|+\mathrm{i}x_1]/p}}{1 - e^{-2\pi[|x_2-y|+\mathrm{i}x_1]/p}} \\
& K_{pq}(\alpha)\frac{p}{2\pi}\ln\left(1 - e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}\right) \\
& - L_{pq}(\alpha)\,|x_2|\,\frac{e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}}{1 - e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}} \Bigg\} \\
& + \frac{\mathrm{i}}{\pi}M_{pq}(-\beta)\left(\frac{1}{2}\sum_{n=-\infty}^{\infty}\mathrm{sgn}[\beta(y-x_2) - x_1 + w - np] - \frac{\beta(y-x_2) - x_1 + w}{p}\right) \\
& - \frac{\mathrm{i}}{\pi}M_{pq}(\alpha)\left(\frac{1}{2}\sum_{n=-\infty}^{\infty}\mathrm{sgn}[\alpha(x_2-y) - x_1 - np] - \frac{\alpha(x_2-y) - x_1}{p}\right). \qquad (6.21)
\end{aligned}
$$

The stresses due to the array of inclusions are now defined completely.

6.3 The Case of Cubic Materials

The stresses due to a periodic array of two-dimensional inclusions, given by (6.8), are given in the previous section for a general uniform lattice mismatch. In this section we concentrate on the special case of the two materials being cubic, so that the mismatch strain e_{ij}^m is given by (6.2) with

$$\epsilon_{ij}^m = f_m \delta_{ij}, \qquad (6.22)$$

where f_m is the fractional difference between the lattice constants of the two materials:

$$f_m = \frac{a_i - a_s}{a_s},$$

where a_i is the lattice constant of the inclusion and a_s is the lattice constant of the surrounding material. This special case is of relevance to the semiconductor industry, where the materials

used are cubic, and where the growth of lattice-mismatched epitaxial layers has gained great importance. With ϵ_{ij}^m given by (6.22), the coefficients K_{pq}, L_{pq}, and M_{pq} (defined by equation (6.17)) take very simple forms: using the symbolic manipulation programme REDUCE it is found that the coefficients L_{pq} are all zero, and the only non-zero coefficients are

$$K_{11}(\gamma) = 2\mu\left(\frac{1+\nu}{1-\nu}\right) f_m \frac{\pi}{\gamma - \mathrm{i}}, \quad K_{12}(\gamma) = -\mathrm{i}K_{11}(\gamma), \quad K_{22}(\gamma) = -K_{11}(\gamma),$$

$$M_{11}(\gamma) = -\frac{2\mu f_m \pi \mathrm{i}(1+\nu)(\gamma^2\nu + 1 - \nu)}{(1-2\nu)(1-\nu)(\gamma^2+1)}, \quad M_{12}(\gamma) = \frac{2\mu f_m \pi \mathrm{i}(1+\nu)\gamma}{(1-\nu)(\gamma^2+1)},$$

$$M_{22}(\gamma) = -\frac{2\mu f_m \pi \mathrm{i}(1+\nu)(\nu + \gamma^2(1-\nu))}{(1-2\nu)(1-\nu)(\gamma^2+1)}, \quad M_{33}(\gamma) = -\frac{2\mu f_m \pi \mathrm{i}\nu(1+\nu)}{(1-2\nu)(1-\nu)}.$$

The stresses due to a periodic array of inclusions, depicted in Fig. 1, may be determined through equations (6.8), (6.9), (6.10), (6.20), and (6.21). For $x_2 > y$ the non-zero stresses are:

$$\begin{aligned}\sigma_{11}(\mathbf{x}) = -\frac{\mu}{\pi}\left(\frac{1+\nu}{1-\nu}\right) f_m \mathrm{Re}\ \Big\{ & -\frac{1}{\beta+\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2| + \mathrm{i}(x_1 - w - \beta y)]/p}\right) \\ & + \frac{1}{\beta+\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2 - y| + \mathrm{i}(x_1 - w)]/p}\right) \\ & + \frac{1}{\alpha-\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2 - y| + \mathrm{i}x_1]/p}\right) \\ & - \frac{1}{\alpha-\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2| + \mathrm{i}(x_1 + \alpha y)]/p}\right) \Big\},\end{aligned}$$

$$\begin{aligned}\sigma_{12}(\mathbf{x}) = -\frac{\mu}{\pi}\left(\frac{1+\nu}{1-\nu}\right) f_m \mathrm{Im}\ \Big\{ & -\frac{1}{\beta+\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2| + \mathrm{i}(x_1 - w - \beta y)]/p}\right) \\ & + \frac{1}{\beta+\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2 - y| + \mathrm{i}(x_1 - w)]/p}\right) \\ & + \frac{1}{\alpha-\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2 - y| + \mathrm{i}x_1]/p}\right) \\ & - \frac{1}{\alpha-\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2| + \mathrm{i}(x_1 + \alpha y)]/p}\right) \Big\},\end{aligned}$$

and

$$\begin{aligned}\sigma_{22}(\mathbf{x}) = \frac{\mu}{\pi}\left(\frac{1+\nu}{1-\nu}\right) f_m \mathrm{Re}\ \Big\{ & -\frac{1}{\beta+\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2| + \mathrm{i}(x_1 - w - \beta y)]/p}\right) \\ & + \frac{1}{\beta+\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2 - y| + \mathrm{i}(x_1 - w)]/p}\right) \\ & + \frac{1}{\alpha-\mathrm{i}} \ln\left(1 - e^{-2\pi[|x_2 - y| + \mathrm{i}x_1]/p}\right)\end{aligned}$$

$$-\frac{1}{\alpha-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}\right)\Bigg\}. \qquad (6.23)$$

For $x_2 < 0$

$$\begin{aligned}\sigma_{11}(\mathbf{x}) = -\frac{\mu}{\pi}\left(\frac{1+\nu}{1-\nu}\right) f_m \mathrm{Re}\ \Bigg\{ & -\frac{1}{\beta-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1-w-\beta y)]/p}\right) \\ & +\frac{1}{\beta-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}(x_1-w)]/p}\right) \\ & +\frac{1}{\alpha+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}x_1]/p}\right) \\ & -\frac{1}{\alpha+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}\right)\Bigg\},\end{aligned}$$

$$\begin{aligned}\sigma_{12}(\mathbf{x}) = \frac{\mu}{\pi}\left(\frac{1+\nu}{1-\nu}\right) f_m \mathrm{Im}\ \Bigg\{ & -\frac{1}{\beta-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1-w-\beta y)]/p}\right) \\ & +\frac{1}{\beta-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}(x_1-w)]/p}\right) \\ & +\frac{1}{\alpha+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}x_1]/p}\right) \\ & -\frac{1}{\alpha+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}\right)\Bigg\},\end{aligned}$$

and

$$\begin{aligned}\sigma_{22}(\mathbf{x}) = \frac{\mu}{\pi}\left(\frac{1+\nu}{1-\nu}\right) f_m \mathrm{Re}\ \Bigg\{ & -\frac{1}{\beta-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1-w-\beta y)]/p}\right) \\ & +\frac{1}{\beta-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}(x_1-w)]/p}\right) \\ & +\frac{1}{\alpha+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}x_1]/p}\right) \\ & -\frac{1}{\alpha+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}\right)\Bigg\}. \qquad (6.24)\end{aligned}$$

For $0 < x_2 < y$

$$\begin{aligned}\sigma_{11}(\mathbf{x}) \quad = \quad & -\frac{\mu}{\pi}\left(\frac{1+\nu}{1-\nu}\right) f_m \mathrm{Re}\Bigg\{-\frac{1}{\beta+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1-w-\beta y)]/p}\right) \\ & +\frac{1}{\beta-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}(x_1-w)]/p}\right) \\ & +\frac{1}{\alpha+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}x_1]/p}\right)\end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{\alpha-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}\right)\Bigg\}\\
&+2\mu\left(\frac{1+\nu}{1-\nu}\right)f_m\frac{\beta^2}{\beta^2+1}\left(\frac{1}{2}\sum_{n=-\infty}^{\infty}\mathrm{sgn}[x_1-\beta(y-x_2)-w-np]-\frac{x_1-\beta(y-x_2)-w}{p}\right)\\
&-2\mu\left(\frac{1+\nu}{1-\nu}\right)f_m\frac{\alpha^2}{\alpha^2+1}\left(\frac{1}{2}\sum_{n=-\infty}^{\infty}\mathrm{sgn}[x_1+\alpha(y-x_2)-np]-\frac{x_1+\alpha(y-x_2)}{p}\right)\\
&-2\mu\left(\frac{1+\nu}{1-\nu}\right)f_m\frac{(\alpha+\beta)(y-x_2)+w}{p},
\end{aligned}
$$

$$
\begin{aligned}
\sigma_{12}(\mathbf{x}) \;=\;& \frac{\mu}{\pi}\left(\frac{1+\nu}{1-\nu}\right)f_m\mathrm{Im}\left\{\frac{1}{\beta+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1-w-\beta y)]/p}\right)\right.\\
&+\frac{1}{\beta-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}(x_1-w)]/p}\right)\\
&+\frac{1}{\alpha+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}x_1]/p}\right)\\
&\left.+\frac{1}{\alpha-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}\right)\right\}\\
&-2\mu\left(\frac{1+\nu}{1-\nu}\right)f_m\frac{\beta}{\beta^2+1}\left(\frac{1}{2}\sum_{n=-\infty}^{\infty}\mathrm{sgn}[x_1-\beta(y-x_2)-w-np]-\frac{x_1-\beta(y-x_2)-w}{p}\right)\\
&-2\mu\left(\frac{1+\nu}{1-\nu}\right)f_m\frac{\alpha}{\alpha^2+1}\left(\frac{1}{2}\sum_{n=-\infty}^{\infty}\mathrm{sgn}[x_1+\alpha(y-x_2)-np]-\frac{x_1+\alpha(y-x_2)}{p}\right),
\end{aligned}
$$

$$
\begin{aligned}
\sigma_{22}(\mathbf{x}) \;=\;& \frac{\mu}{\pi}\left(\frac{1+\nu}{1-\nu}\right)f_m\mathrm{Re}\left\{-\frac{1}{\beta+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1-w-\beta y)]/p}\right)\right.\\
&+\frac{1}{\beta-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}(x_1-w)]/p}\right)\\
&+\frac{1}{\alpha+\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2-y|+\mathrm{i}x_1]/p}\right)\\
&\left.-\frac{1}{\alpha-\mathrm{i}}\ln\left(1-e^{-2\pi[|x_2|+\mathrm{i}(x_1+\alpha y)]/p}\right)\right\}\\
&+2\mu\left(\frac{1+\nu}{1-\nu}\right)f_m\frac{1}{\beta^2+1}\left(\frac{1}{2}\sum_{n=-\infty}^{\infty}\mathrm{sgn}[x_1-\beta(y-x_2)-w-np]-\frac{x_1-\beta(y-x_2)-w}{p}\right)\\
&-2\mu\left(\frac{1+\nu}{1-\nu}\right)f_m\frac{1}{\alpha^2+1}\left(\frac{1}{2}\sum_{n=-\infty}^{\infty}\mathrm{sgn}[x_1+\alpha(y-x_2)-np]-\frac{x_1+\alpha(y-x_2)}{p}\right),
\end{aligned}
$$

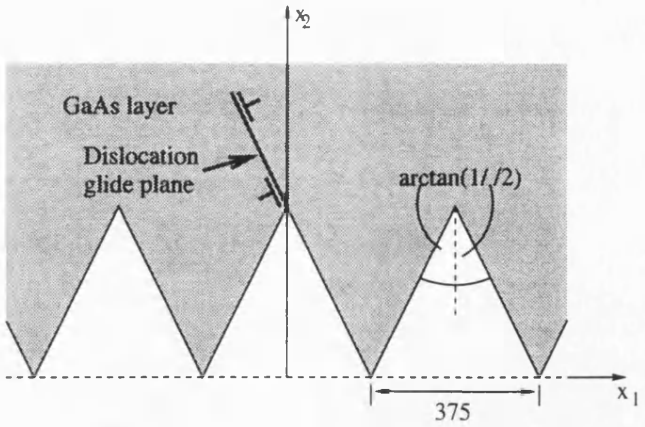

Figure 6.3: Schematic illustration of the GaAs layer grown on a saw-tooth patterned Si(001) substrate by Sprung *et al.* Also illustrated is a dislocation loop nucleating on a {111} glide plane at the apex of one of the 'teeth'. Lengths are in nanometres.

and

$$\sigma_{33}(\mathbf{x}) = -\mu\left(\frac{1+\nu}{1-\nu}\right) f_m \sum_{n=-\infty}^{\infty} \left(\mathrm{sgn}[x_1 + \alpha(y-x_2) - np] - \mathrm{sgn}[x_1 - \beta(y-x_2) - w - np]\right). \tag{6.25}$$

Note that the function

$$\sum_{n=-\infty}^{\infty} \mathrm{sgn}(z-np) = \lim_{N\to\infty} \sum_{n=-N}^{N} \mathrm{sgn}(z-np)$$

simply represents sgn$z$ plus twice the integer part of z/p.

6.4 Structural Stability of Layers Grown on Patterned or Off-cut Substrates

As an illustration of how the formulae obtained here may be used to assess the structural stability of growth on patterned substrates we consider the GaAs layer grown on a Si(001) saw-tooth-patterned substrate by Sprung *et al*[62]. The structure is illustrated in Fig. 5.4. By growing GaAs on saw-tooth-patterned Si they were able to exhibit single domain growth and good crystal structure using a nominally exact (001) Si wafer. The motivation of their study was the industrial desirability of using standard Si(001) wafers for GaAs layer growth. The stresses and strains in the system are obtained by considering the configuration of Fig. 5.4 as

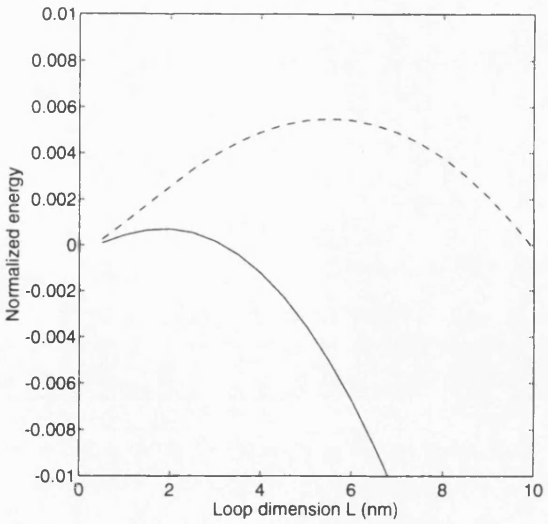

Figure 6.4: Energy, normalized by μb^3, as a function of side length L for a square 60° glide loop nucleating at a saw-tooth patterned substrate-layer interface in the manner illustrated in Fig. 5.4 (solid line) and at a planar (001) substrate-layer interface (broken line).

the superposition of a uniformly strained layer lying in the region $x_2 > 0$ and a periodic array of triangular inclusions of negative mismatch based at $x_2 = 0$. Taking the lattice constants of GaAs and Si to be 5.65 and 5.43Å respectively, the fractional lattice mismatch is $f_m = 0.041$.

The apex of one of the 'teeth' of the grating is an obvious candidate for a prefered dislocation nucleation site. To investigate this we consider the formation of a square 60° glide loop on a $\{111\}$ plane that meets one of the teeth of the grating at its apex, as illustrated in Fig. 5.4. A typical Burgers vector of such a dislocation is

$$\mathbf{b} = b(-1/2, 1/\sqrt{2}, -1/2),$$

where b, the magnitude of the Burgers vector, is taken to be 4Å. The self-energy of the loop may be written obtained from (5.11). Because the solution for the stresses due to the non-planar interface features does not account for the free surface effects, we use just the infinite-body value for the self-energy of the dislocation in the calculations described here. The interaction energy W_i, given in (5.5), is evaluated numerically, using the stresses given by (5.15), (6.23), (6.24), and (6.25). The energy variations with loop dimension for nucleation at the apex of a sawtooth and for nucleation at a planar interface are given by, respectively, the solid and dashed lines in Fig. 5.5. The energies have been normalized by μb^3. The activation energy, which is the maximum loop energy, is lower by a factor five for nucleation at a sawtooth as compared with nucleation

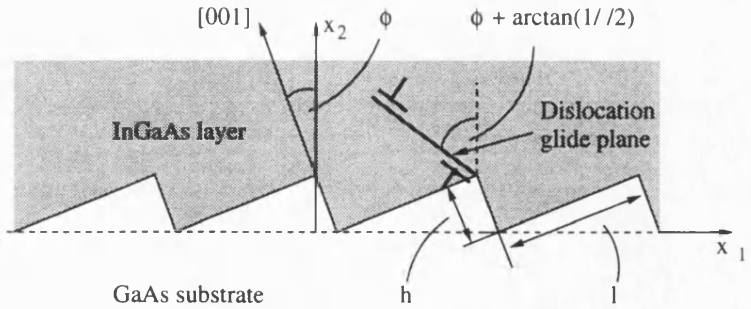

Figure 6.5: Schematic illustration of an InGaAs strained layer grown on a GaAs substrate off-cut by an angle ϕ from (001) towards (111). Also illustrated is a dislocation loop nucleating on a $\{111\}$ glide plane at an interface step.

at a planar interface. Such a lowering in the activation barrier to nucleation would be expected to yield an enormous increase in the anticipated nucleation rate, obtained from (5.1).

A similar, although less extreme, lowering of the activation barrier to loop nucleation can be shown to occur due to substrate off-cutting. The interface between a layer and substrate, where the substrate is off-cut by an angle of ϕ from (001) towards (111), is illustrated in Fig. 5.6. The off-cut causes the formation of interface ledges of length l ending in steps that are of height h, where $h/l = \tan\phi$. For the purposes of stress evaluation, the decomposition of the configuration illustrated in Fig. 5.6, into a strained layer with a planar substrate-layer interface and a periodic array of triangular inclusions, is identical with that for the saw-tooth interface just discussed. It has been shown that 'step-bunching' tends to occur[1, 73], so that the steps are not of monolayer height, but may be 2-8 monolayers high. We consider the nucleation of a square 60° dislocation at the apex of a step, on the $\{111\}$ glide plane illustrated in Fig. 5.6. Because of the off-cut, the glide plane is inclined at an angle $\phi + \arctan(1/\sqrt{2})$ to the interface normal. Results for the activation barrier to loop nucleation in $In_xGa_{1-x}As$ layers grown on a GaAs substrate off-cut by 2° from (001) towards (110) are given in Fig. 5.7, for values of x ranging from 0 to 1, normalized to the corresponding activation energies for nucleation at a flat interface. The step height was taken to be 2nm, corresponding to 5-10 monolayers. The activation energy is reduced by the presence of the step by up to 30-40%. As would be expected, the reduction is not as severe as for the case of a patterned substrate, because the scale of the interface non-uniformities is smaller. However, since the nucleation rate may be expected to depend exponentially on the activation energy, via an Arrhenius-type term, the effect, on predicted nucleation rates, could still be significant. Hence interface steps may be expected to be favoured sites for dislocation nucleation, as some experimental observations have suggested[57].

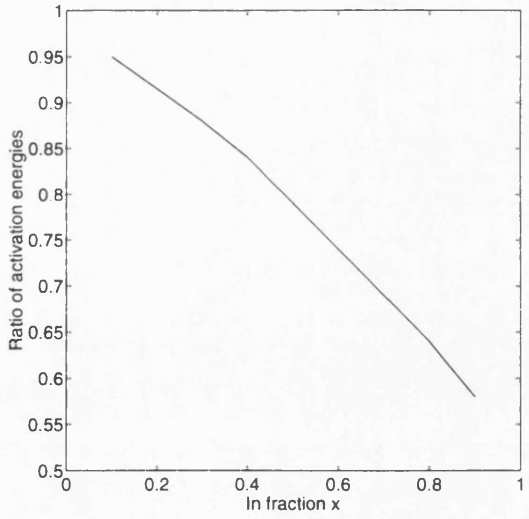

Figure 6.6: Ratio of activation energy for formation of a 60° square loop at an interface step, as illustrated in Fig. 5.6, to that for formation of a 60° square loop at a planar (001) interface, plotted as a function of In fraction x. The substrate was taken to be off-cut by 2° from (001) towards (111).

6.5 The Electronic Properties of Quantum Wires

Although this thesis is mainly concerned with the mechanical stability of strained semiconductor structures, it is worth presenting a brief discussion of an important application of the solution obtained here in the realm of the assessment of the electronic properties of strained quantum wire arrays. A buried strained quantum wire is simply a two-dimensional lattice-mismatched inclusion. A quantum wire array is simply a periodic arrangement of such inclusions, of precisely the type illustrated in Fig. 6.2. The strains in such a configuration have been derived in the preceding sections. Strained quantum wire arrays are increasingly being exploited in experimental device designs, and it is therefore of great interest to attempt an assessment of the effects of the strain state in such a structure on its electronic properties.

In InGaAs, strain causes uniform shifting of the conductance and valence bands, and splitting of the light and heavy hole valence bands. Under compressive strain the heavy hole band moves up, so it is the energy gap between this and the conduction band that defines the band gap[8]. Under these conditions the change in band gap is given by[8]

$$\Delta E_g = a\epsilon'_{kk} - \sqrt{\Sigma_\epsilon}, \tag{6.26}$$

where

$$\Sigma_\epsilon = \sum_{i<j} \left[\frac{b^2}{2} (\epsilon'_{jj} - \epsilon'_{ii})^2 + d^2 (\epsilon'_{ij})^2 \right]. \tag{6.27}$$

Here ϵ'_{ij} represents the *total* strains (corresponding to $e_{ij} - e^m_{ij}$, which are obtained by inverting (6.1)) in the crystallographic coordinate system, with axes chosen to coincide with the $\langle 100 \rangle$ directions. In the above, a is the *interband* dilatational (hydrostatic) deformation potential[58]. The deformation potentials b and d follow the notation of Bir and Pikus[8]. In InGaAs the net effect of compressive strain is to increase the band gap by an amount determined almost entirely by the dilatational component of the strain: the valence band splitting is relatively insignificant. Now, it may be seen explicitly from (6.23), (6.24), and (6.25) that the invariant dilatation

$$\epsilon'_{kk} = \frac{(1-2\nu)\sigma_{kk}}{2\mu(1+\nu)} \tag{6.28}$$

is zero outside the mismatched regions and constant within them with value

$$\epsilon'_{kk} = -2 \left(\frac{1-2\nu}{1-\nu} \right) f_m, \tag{6.29}$$

which is independent of the aspect ratio and spacing of the inclusions. This result holds true for two-dimensional inclusions of any geometry, and may be proved directly from (6.5), as shown in Appendix F. In a quantum well (modelled as a layer of infinite lateral extent) the non-zero stresses are

$$\sigma_{11} = -2\mu \frac{1+\nu}{1-\nu} f_m = \sigma_{33}.$$

It is clear, therefore, that the dilatation in a quantum well is also given by (6.29), the same as for a quantum wire. Since the band gap shift in InGaAs is dominated by the dilatational component, very little difference would be expected between the band gaps of quantum wells and wires of the same composition in this material system. This fact has been verified experimentally[2].

The effect of strain on the band gap of GeSi is somewhat different from that described above for InGaAs. In GeSi the six-fold degenerate conduction band splits, tending to reduce the band gap. Equation (6.26) still describes the band gap narrowing due to the dilatation and the splitting of the valence band, but an additional term is required to describe the effect of the splitting of the conduction band. The total band gap narrowing is described for GeSi by[58, 70]

$$\Delta E_g = a\epsilon'_{kk} - \sqrt{\Sigma_\epsilon} + \Xi_u \min_i \{\epsilon'_{ii}\}, \tag{6.30}$$

where $\min\{\epsilon'_{ii}\}$ is the most negative ii-strain, since it is this strain that determines the energy of

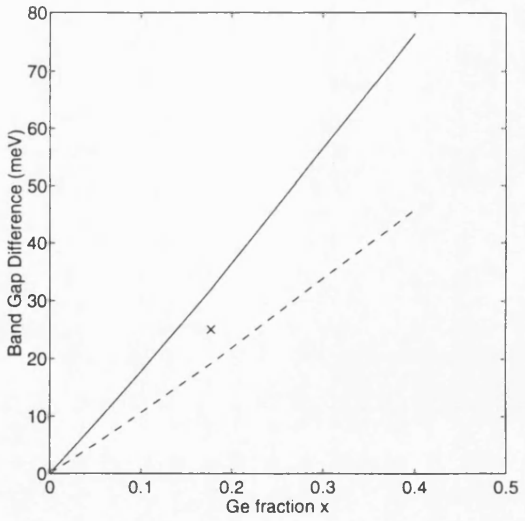

Figure 6.7: Predicted difference between the band gaps of Ge_xSi_{1-x} triangular quantum wires, buried in Si, and quantum wells of the same composition, plotted as a function of Ge fraction x for $p/w = 1$ (dashed line) and $p/w = 10$ (solid line). The cross marks the value for the difference obtained experimentally for $p/w = 10$.

the lowest conduction band minimum, thereby defining the band gap. The effect of conduction band splitting outweighs the dilatational band gap widening, and the net effect of compressive strain is to reduce the band gap of GeSi[70]. In this case the difference between the band gaps of quantum wells and wires of the same composition can be significant. In Fig. 6.7 we plot the difference in band gap between Ge_xSi_{1-x} triangular quantum wires (embedded in Si) and quantum wells of the same composition as a function of the Ge fraction x for two values of the ratio p/w, where w is the width of the base of a wire and p is the period of the array. The wires were taken parallel to the [110] direction and to lie in a (001) plane, and the normal to the quantum well was taken to be [001]. The triangular wires were chosen to be bounded by planes in the (001), $(\bar{1}11)$, and $(1\bar{1}1)$ orientations. A coordinate transfomation was required to obtain the strains in the crystallographic coordinate system. Of course the strains are not constant over the wires. We have taken the average of each strain component over the wire, and used these values in (6.30). We took the values $a = -1.37$eV, $b = -2.10$eV, $d = -4.85$eV, and $\Xi_u = 8.6$eV for Si[45], and $a = -1.83$eV[70], $b = -2.86$eV[6], $d = -5.28$eV[11], and $\Xi_u = 9.42$eV[70] for Ge, using linear interpolation to obtain deformation potentials for the alloys. Poisson's ratio was taken to be 0.25. The results for the case $p/w = 10$ are almost indistinguishable from those for a single quantum wire. The difference between the band gaps of the two structures is of the order of several tens of meV, which is not insignificant.

The results of photoluminescence spectroscopy are available in the literature[68] for quantum wires and a quantum well of composition $Ge_{0.177}Si_{0.823}$. The quantum wires were triangular and arranged in arrays with spacing corresponding approximately to $p/w = 10$. The wires were in the [110] direction, and were bounded by planes in the (001), $(\bar{1}11)$, and $(1\bar{1}1)$ orientations, as for the above calculations. The spectra revealed a clear difference between the band gaps of the two structures, the band gap of the quantum wires being approximately 25meV larger than that of the quantum well. This value is marked by the cross on Fig. 6.7, and is in reasonable agreement with the theoretical value of 32meV obtained here for triangular quantum wires with $p/w = 10$. Given the uncertainties involved in the rudimentary band gap calculation presented here (formulae for uniform strains have been used whereas the strains vary within the wires, and no allowance for quantum size effects has been made) the agreement obtained here should be regarded as quite satisfactory.

In summary, the formulae derived in this chapter for the stresses due to a periodic array of mismatched inclusions have been used to demonstrate that there should be little difference between the band gaps of InGaAs quantum wells and quantum wires. By contrast, the band gap of a GeSi quantum wire may differ from that of the corresponding quantum well by several tens of meV. The predicted difference is in good agreement with experimental observation.

Chapter 7

Closure

The aim of this thesis has been to make progress in the area of predicting the nature and extent of mechanical degradation in strained semiconductor devices. In pursuing this aim, three solutions have been obtained that should have applications in many areas, not just within the narrow remit of the investigation presented here. In the context of the present work these solutions have been used to gain insight, both qualitative and quantitative, into the strain relaxation process.

The solution for arrays of dislocations that was derived in Chapter 2 has been used in Chapter 3 to study the propagation via threading motion of dislocations in a strained layer, taking into account interactions between parallel and perpendicular dislocations. The incorporation of dislocation interactions in such detail with the acknowledgement of elastic anisotropy represent advances on previous work. The sequential view of strain relaxation, which is based on consideration of the motion of a dislocation via threading motion, was shown to yield better agreement with experiment than a simple energy minimization approach, which takes no account of the method of dislocation propagation. However, even the sequential approach was found to overestimate by a large amount the extent of strain relaxation. This discrepency is thought to arise in part from the approximation of perfect periodicity in the dislocation distribution, and in part through the lack of any kinetc features in the model.

An important kinetic feature is the rate at which dislocations nucleate. Study of the nucleation process was considered in Chapter 5. The availability of the line-integral representation for the stresses due to an arbitrary dislocation in a half space, which was derived in Chapter 4, enabled estimates to be obtained for dislocation activation energies taking the effects of a free surface into account. Previous approximations, based on infinite-body results, were found

to lead to large errors. The improved estimates for activation energies were used to predict the lattice mismatch at which homogeneous nucleation should become experimentally significant. It was found that the same lattice mismatch also marks the threshold at which a mechanism for the conversion of 60° dislocations into 90° dislocations should switch on. The commencement of these two occurences was used to explain the observed change in the type and distribution of dislocations as the lattice mismatch is increased.

The third new solution, for the stresses due to an array of two-dimensional mismatched inclusions in an elastic body, has not yet been used to its full potential. However, it was illustrated how the solution can aid the study of nucleation at interface steps and other non-planar interface features. Moreover, the solution was used to explain the observed differences between the strain-induced band gap shifts in GeSi/Si quantum wires and quantum wells of the same composition and the lack of such differences in the InGaAs/GaAs system. Since previous band-structure calculations appear to have been hampered by the lack of a correct solution for the stresses in a quantum wire array[72, 76, 77], this solution will be of great use in the assessment of the electronic properties of such a structure.

Although the processes of dislocation propagation and nucleation have been considered here separately in some detail, no real attempt has been made to combine the two processes in an overall kinetic model of strain relaxation. The development of such a model, drawing on the results obtained here, is an important project for the future. To illustrate the problems that would need to be addressed in the development of such a model, a rudimentary outline structure is now given. The problem may be defined in terms of the total length, L, of misfit segment at the interface, which is related to the number of mobile threading dislocations, N_m, and the mean threading segment velocity, $\overline{v}$, via the following differential equation:

$$\frac{dL}{dt} = N_m \overline{v}. \tag{7.1}$$

The number of mobile threading dislocations evolves according to

$$\frac{dN_m}{dt} = -J_{block} + J_{nuc} + J_{mult}, \tag{7.2}$$

where J_{block} is the rate at which threading dislocations are rendered immobile through being blocked by a perpendicular dislocation, and J_{nuc} and J_{mult} are the rates at which new mobile threading dislocations are created by nucleation and multiplication events, respectively. Assuming that the dislocation distribution evolves so that long-range shear strains are annuled, the

relevant observable may be taken to be the mean in-plane strain in the layer, given by

$$\epsilon = -f_m - \frac{b_1 L}{2A}, \tag{7.3}$$

where A is the surface area of the layer, and b_1 is the strain-relieving compoinent of the Burgers vector, which is of opposite sign to f_m. Using (7.3) we may write (7.1) as

$$\frac{d\epsilon}{dt} = -\frac{b_1 N_m \overline{v}}{2A}. \tag{7.4}$$

Given suitable expressions for $\overline{v}$, J_{block}, J_{nuc} and J_{mult} in terms of ϵ and N_m, (7.2) and (7.4) provide a pair of first order differential equations that may be solved with ease numerically.

Theoretical and empirical expressions for the velocity of a threading segment are of the form[31, 35, 49, 66]

$$v = \frac{MG}{bh}, \tag{7.5}$$

where G is the driving force on the dislocation, given by (3.5), and M is a mobility factor. Averaging (7.5) with the aid of (3.5) yields

$$\overline{v} = -\frac{ME^s}{bh} - M(b_i/b)m_j\Gamma_{ijkl}e^r_{kl}. \tag{7.6}$$

Of course, the dislocations are of finite length, so an interpretation of e^r_{kl} is required. Motivated by (7.3) and (2.47),f e^r_{kl} is defined by

$$e^r_{kl} = f_m\delta_{kl} + \frac{b_1 L}{2A}(\delta_{k1}\delta_{l1} + \delta_{k3}\delta_{l3}).$$

The quantity $L/2A$ plays the role of an effective reciprocal dislocation spacing, $1/p$. The terms e^r_{k3}, $k \neq 3$ have been chosen to be zero, it being asumed that long-range shear strains are annuled. The choice of e^r_{k2}, $k = 1, 2, 3$, is not of importance, since $\Gamma_{ijk2} = 0$. Equation (7.4) is now defined completely.

We now require expressions for J_{block}, J_{nuc}, and J_{mult}. If $N_i(t)$ denotes the *total* number of intersections of perpendicular dislocations that have occured up to time t, then the blocking rate is given by

$$J_{block} = \frac{dN_i(t)}{dt}P(t). \tag{7.7}$$

Here $P(t)$ is the probability of blocking, given an interaction at time t, which is equal to $1/4$ if $p_a > \overline{p} \equiv 2A/L$ (where p_a was defined in Section 3.3 to be the dislocation spacing for which

blocking commences) and is equal to 0 otherwise. The factor 1/4 appears because only one quarter of interactions between perpendicular dislocations are repulsive, leading to blocking. If dislocations were strongly clustered then the equations of Chapter 3 would have to be modified to take this into account. The most elementary modification would be to multiply the interaction force between dislocations by the average number of dislocations in a cluster. A rudimentary estimate of N_i, the mean number of intersections, is given by

$$N_i = \frac{L^2}{4A}, \tag{7.8}$$

which may be shown to be rigorously correct if the dislocations are all of the same length and randomly distributed, and which should be accurate enough, at least for a preliminary investigation. Differentiation of (7.8) yields, in combination with (7.1), an expression for dN_i/dt. Equations (7.7) and (7.8), together with condition (a) of Chapter 3, provide an approximate description of the rate at which dislocations are rendered immobile through blocking.

It is not easy to write even such rudimentary equations to govern the nucleation and multiplication rates, J_{nuc} and J_{mult}. The contribution from homogeneous nucleation can be estimated using techniques similar to those employed in Section 5.3, with the surface area available for nucleation being modelled as some function of the total length of dislocation, L. However, the contribution from heterogenous nucleation and multiplication would seem to be very difficult to assess. Not only would rates have to be calculated for nucleation at heterogeneous nucleation sites, such as the interface features considered in Chapter 6, but experimental evidence for the distribution of such inhomgeneities would have to be obtained. Given the large number of possible inhomogeneities, this task would be extremely difficult. Similarly, modelling of dislocation multiplication s required at a fundamental level. For example, consider the Hagen-Strunk mechanism[30]. This mechanism operates when two perpendicular dislocations intersect that have parallel Burgers vectors. Annihilation occurs at the intersection, producing two L-shaped dislocations. One of the corners is then repelled from the interface towards the free surface, generating, on meeting the free surface, precisely two new mobile threading dislocations. Because there is a one in four chance of the Burgers vectors of two perpendicular dislocations being parallel, the multiplication rate for this mechanism is given by

$$J_{mult} = \frac{1}{2}\frac{dN_i}{dt}P_{mult},$$

where the term P_{mult} is the probability that an interaction of the appropriate type leads to a multiplication event: this would be expected to be, for example, mismatch and thickness

dependent. As illustrated by Lefebvre *et al.*[46], unless the mismatch strain is rather small, wht repulsive interaction between the corners is not enough to drive one of the corners to the free surface, especially in thick layers. In fact, evidence seems to suggest that the Hagen-Strunk mechanism does not operate to any great extent in the InGaAs/GaAs or GeSi/Si systems[47, 46], and that instead regenerative multiplication sources of a complex nature operate. These sources also occur when dislocations with parallel Burgers vectors intersect, but each intersection event produce mores than one dislocation at a rate that will be determined by an activation energy in a manner similar to a normal nucleation event. As dislocations are created by multiplication at each intersection, the stress fields that they set up will eventually prevent further multiplication occuring there. The number of dislocations generated by a multiplication source, for a given mismatch and layer thickness, before deactivation occurs is an important parameter to determine. Unfortunately, the complex nature of the multiplication mechanisms that appear to be operating[47, 46] will make a theoretical estimate of the multiplication activation energy rather hard to obtain, but such an estimate would be necessary for a complete theoretical description. An alternative approach would be to use the multiplication activation energy as a fitting parameter. For example, LeGoues *et al.*[47] ingeniously used properties of systems grown on tilted substrates to determine experimentally the multiplication activation energy in graded buffer layers. The disadvantage of such an approach is that the time (and hence strain) *average* of the multiplication rate is measured, rather than the strain dependence. This limitation makes doubtful the applicability of the results to experiments with conditions different from those under which the results were obtained. However, a combination of careful qualitative modelling and quantitative experimental measurement along the lines of the method of LeGoues *et al.* seems to be a promising approach.

In this final Chapter we have illustrated an approach that might be taken to developing a model of the strain relaxation process that interconnects the elements of dislocation propagation, interaction, nucleation, and multiplication. In the course of outlining this approach it has become apparent that much work remains to be done before a reliable predictive model can be constructed. However, it is hoped that some of this work is made easier in the light of the studies presented here of dislocation interactions and nucleation.

Appendix A

We shall verify the following identities:

$$\sum_r \left[\frac{\mathbf{N}_r (\mathbf{C}_r^{(2)})^{\mathrm{T}}}{\frac{\partial D}{\partial \xi_2}(1, \gamma_r)} + \overline{\frac{\mathbf{N}_r (\mathbf{C}_r^{(2)})^{\mathrm{T}}}{\frac{\partial D}{\partial \xi_2}(1, \overline{\gamma_r})}} \right] = \mathbf{I}, \tag{A.1}$$

$$\sum_r \left[\frac{\mathbf{C}_r^{(2)} \mathbf{N}_r (\mathbf{C}_r^{(2)})^{\mathrm{T}}}{\frac{\partial D}{\partial \xi_2}(1, \gamma_r)} + \overline{\frac{\mathbf{C}_r^{(2)} \mathbf{N}_r (\mathbf{C}_r^{(2)})^{\mathrm{T}}}{\frac{\partial D}{\partial \xi_2}(1, \overline{\gamma_r})}} \right] = 0, \tag{A.2}$$

$$\sum_r \left[\frac{\mathbf{N}_r}{\frac{\partial D}{\partial \xi_2}(1, \gamma_r)} + \overline{\frac{\mathbf{N}_r}{\frac{\partial D}{\partial \xi_2}(1, \overline{\gamma_r})}} \right] = 0, \tag{A.3}$$

and

$$\sum_r \left[\frac{\mathbf{C}_r^{(2)} \mathbf{N}_r}{\frac{\partial D}{\partial \xi_2}(1, \gamma_r)} + \overline{\frac{\mathbf{C}_r \mathbf{N}_r}{\frac{\partial D}{\partial \xi_2}(1, \overline{\gamma_r})}} \right] = \mathbf{I}. \tag{A.4}$$

Here $\mathbf{I}$ represents the identity matrix. Writing $\boldsymbol{\xi} = (\xi_1, \xi_2)$, consideration of (2.19), (2.21), and (2.25) reveals that $\mathbf{L}^{-1}(\boldsymbol{\xi})$, $\mathbf{C}^{(2)}(\boldsymbol{\xi})$, and $D(\boldsymbol{\xi})$ are homogeneous functions of $\boldsymbol{\xi}$ of degrees -2, 1, and 6 respectively. Further, it may be demonstrated that

$$\left.\begin{aligned} \mathbf{L}^{-1}(\mathbf{C}^{(2)})^{\mathrm{T}}(\boldsymbol{\xi}) &\sim \tfrac{\mathbf{I}}{|\xi_2|}, \\ \mathbf{C}^{(2)}\mathbf{L}^{-1}(\mathbf{C}^{(2)})^{\mathrm{T}}(\xi) - \mathbf{C}^{(2)}(0,1) &\sim \bigcirc\left(\tfrac{1}{|\xi_2|^2}\right), \\ \mathbf{L}^{-1}(\boldsymbol{\xi}) &\sim \bigcirc\left(\tfrac{1}{|\xi_2|^2}\right), \\ \mathbf{C}^{(2)}\mathbf{L}^{-1}(\boldsymbol{\xi}) &\sim \tfrac{\mathbf{I}}{|\xi_2|} \end{aligned}\right\} \text{ as } |\xi_2| \to \infty. \tag{A.5}$$

Now regard ξ_2 as a complex variable and consider integration of each of the quantities $\mathbf{L}^{-1}(\mathbf{C}^{(2)})^T(\boldsymbol{\xi})$, $\mathbf{C}^{(2)}\mathbf{L}^{-1}(\mathbf{C}^{(2)})^{\mathrm{T}}(\xi) - \mathbf{C}^{(2)}(0,1)$, $\mathbf{L}^{-1}(\boldsymbol{\xi})$, and $\mathbf{C}^{(2)}\mathbf{L}^{-1}(\boldsymbol{\xi})$ around a circular contour Γ, large enough to contain the poles of each integrand. Each integral may be evaluated by deforming

the contour Γ to infinity and using the explicit behaviour of the integrands, as given in (A.5), to obtain the result. Alternatively, Cauchy's theorem may be used, after observing that each of the integrands has poles precisely at the roots of $D(\boldsymbol{\xi})$; specifically, at $\xi_2 = \xi_1\gamma_r$, $\xi_2 = \xi_1\overline{\gamma_r}$, r =1, 2, 3. Equating the results of the two methods of integration in each case yields the identities (A.1) through (A.4).

Appendix B

We require the two dimensional Fourier transform of the components of the infinite-body Green's tensor. This is most conveniently obtained from the whole-space transform of the infinite-body Green's tensor, given by

$$\tilde{G}^{\infty}_{mk}(\boldsymbol{\zeta}) = \iiint dx_1 dx_2 dx_3 G^{\infty}_{mk}(\mathbf{x}) e^{\mathrm{i}\boldsymbol{\zeta}.\mathbf{x}}.$$

An expression for $\tilde{G}^{\infty}_{mk}$ is obtained by taking the whole-space transform of (4.5), the defining equation for the Green's tensor, recalling that G^{∞}_{mk} is a function of $\mathbf{x} - \mathbf{x}'$:

$$-\zeta_l \zeta_j c_{ijkl} \tilde{G}^{\infty}_{mk} + \delta_{im} = 0. \tag{B.1}$$

For isotropy, (B.1) yields a direct expression for $\tilde{G}^{\infty}_{mk}$[14]:

$$\tilde{G}^{\infty}_{mk}(\boldsymbol{\zeta}) = \frac{1}{\mu}\left[\frac{\delta_{mk}}{|\boldsymbol{\zeta}|^2} - \left(\frac{\lambda+\mu}{\lambda+2\mu}\right)\frac{\zeta_m \zeta_k}{|\boldsymbol{\zeta}|^4}\right]. \tag{B.2}$$

Now let $\zeta_1 = \xi_1$, $\zeta_2 = \xi_2$, and invert $\tilde{G}^{\infty}_{mk}$ with respect to ζ_3 to obtain

$$\hat{G}^{\infty}_{mk}(\boldsymbol{\xi}, z) = \frac{1}{2\pi}\int d\zeta_3 \tilde{G}^{\infty}_{mk}(\xi_1, \xi_2, \zeta_3) e^{-\mathrm{i}\zeta_3 z}. \tag{B.3}$$

From (B.2) it may be seen that the required integrals take one of the forms

$$\int_{-\infty}^{\infty} d\zeta_3 \frac{e^{-\mathrm{i}\zeta_3 z}}{|\boldsymbol{\xi}|^2 + \zeta_3^2}, \tag{B.4}$$

$$\int_{-\infty}^{\infty} d\zeta_3 \zeta_3^m \frac{e^{-\mathrm{i}\zeta_3 z}}{(|\boldsymbol{\xi}|^2 + \zeta_3^2)^2}, \tag{B.5}$$

where $m = 0$, 1, or 2. The above are required for the two cases $z = x_3 - x_3'$ and $z = x_3 + x_3'$, where $x_3' > x_3 \geq 0$. Therefore, we must consider both positive and negative z. However, it is immediate from (B.3) that

$$\hat{G}^{\infty}_{mk}(\boldsymbol{\xi}, -z) = \overline{\hat{G}^{\infty}_{mk}(\boldsymbol{\xi}, z)}. \tag{B.6}$$

Therefore, we need only consider the case $z < 0$ explicitly. The integrals (B.4) and (B.5) are performed by considering ζ_3 as a complex variable, closing the contour appropriately, and using Cauchy's theorem. Since $z < 0$, the contour must be closed in the upper half plane so as to enclose the pole at $\zeta_3 = \mathrm{i}\,|\boldsymbol{\xi}|$, which is a simple pole in the case of (B.4) and a double pole in the case of (B.5). Straightforward application of Cauchy's residue theorem gives the result that for $z < 0$:

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\infty}^{\infty} d\zeta_3 \frac{e^{-\mathrm{i}\zeta_3 z}}{|\boldsymbol{\xi}|^2 + \zeta_3^2} &= \frac{e^{|\boldsymbol{\xi}| z}}{2\,|\boldsymbol{\xi}|}, \\
\frac{1}{2\pi}\int_{-\infty}^{\infty} d\zeta_3 \frac{e^{-\mathrm{i}\zeta_3 z}}{(|\boldsymbol{\xi}|^2 + \zeta_3^2)^2} &= (1 - |\boldsymbol{\xi}|\, z)\frac{e^{|\boldsymbol{\xi}| z}}{4\,|\boldsymbol{\xi}|^3}, \\
\frac{1}{2\pi}\int_{-\infty}^{\infty} d\zeta_3 \zeta_3 \frac{e^{-\mathrm{i}\zeta_3 z}}{(|\boldsymbol{\xi}|^2 + \zeta_3^2)^2} &= -\mathrm{i}z\frac{e^{|\boldsymbol{\xi}| z}}{4\,|\boldsymbol{\xi}|}, \\
\frac{1}{2\pi}\int_{-\infty}^{\infty} d\zeta_3 \zeta_3^2 \frac{e^{-\mathrm{i}\zeta_3 z}}{(|\boldsymbol{\xi}|^2 + \zeta_3^2)^2} &= (1 + |\boldsymbol{\xi}|\, z)\frac{e^{|\boldsymbol{\xi}| z}}{4\,|\boldsymbol{\xi}|}.
\end{aligned} \tag{B.7}$$

The results of (4.19) are obtained immediately from (B.2), (B.3), and (B.7) by putting $z = x_3 - x_3'$. From (B.6), the additional results required in Section 4.4 are obtained by putting $z = -(x_3 + x_3')$ and taking the complex conjugate of all expressions.

Appendix C

We require the integrals

$$J_Q^{\pm} = \oint_{|\eta|=1} ds \frac{(\eta_1 \pm \mathrm{i}\eta_2)^Q}{x_3 + x_3' + \mathrm{i}\boldsymbol{\eta}.(\mathbf{y} - \mathbf{y}')}.$$

First consider J_Q^+, and make the substitution $\zeta = \eta_1 + \mathrm{i}\eta_2$. Then

$$\begin{aligned} J_Q^+ &= -2\mathrm{i} \oint_{|\zeta|=1} d\zeta \frac{\zeta^Q}{\zeta^2[\mathrm{i}(x_1 - x_1') + (x_2 - x_2')] + 2\zeta(x_3 + x_3') + \mathrm{i}(x_1 - x_1') - (x_2 - x_2')} \\ &= \frac{-2\mathrm{i}}{\mathrm{i}(x_1 - x_1') + (x_2 - x_2')} \oint_{|\zeta|=1} d\zeta \frac{\zeta^Q}{(\zeta - \zeta_+)(\zeta - \zeta_-)}, \end{aligned}$$

where

$$\zeta_{\pm} = \frac{-(x_3 + x_3') \pm \sqrt{(x_1 - x_1')^2 + (x_2 - x_2')^2 + (x_3 + x_3')^2}}{\mathrm{i}(x_1 - x_1') + (x_2 - x_2')}.$$

Only the pole ζ_+ lies within $|\zeta| = 1$, and so by Cauchy's residue theorem

$$J_Q^+ = \frac{4\pi}{\mathrm{i}(x_1 - x_1') + (x_2 - x_2')} \frac{\zeta_+^Q}{\zeta_+ - \zeta_-},$$

which simplifies to the form of equation (4.24). To obtain J_Q^-, make the substitution $\zeta = \eta_1 - \mathrm{i}\eta_2$. Proceeding as for J_Q^+ above, the form of (4.24) is easily obtained.

Appendix D

We require the integrals

$$H_Q^{\pm} = \oint_{|\eta|=1} ds \frac{(\eta_1 \pm \mathrm{i}\eta_2)^Q}{(\alpha_3 - \mathrm{i}\eta_1\alpha_1 - \mathrm{i}\eta_2\alpha_2)[x_3 + a_3 + \mathrm{i}\eta_1(x_1 - a_1) + \mathrm{i}\eta_2(x_2 - a_2)]}.$$

First evaluate H_Q^+ by making the substitution $\zeta = \eta_1 + \mathrm{i}\eta_2$. On making the substitution it is straightforward to demonstrate that

$$H_Q^+ = \frac{4\mathrm{i}}{(\mathrm{i}\alpha_1 + \alpha_2)[\mathrm{i}(x_1 - a_1) + (x_2 - a_2)]} \oint_{|\zeta|=1} d\zeta \frac{\zeta^{Q+1}}{(\zeta - f_+)(\zeta - f_-)(\zeta - g_+)(\zeta - g_-)},$$

where

$$f_{\pm} = \frac{\alpha_3 \pm \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}}{\mathrm{i}\alpha_1 + \alpha_2},$$

$$g_{\pm} = \frac{-(x_3 + a_3) \pm \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + (x_3 + a_3)^2}}{\mathrm{i}(x_1 - a_1) + (x_2 - a_2)}.$$

Only the poles at f_- and g_+ lie within $|\zeta| = 1$, and they are both simple, so application of Cauchy's residue theorem immediately yields the result of (4.28). H_Q^- is evaluated by making the substitution $\zeta = \eta_1 - \mathrm{i}\eta_2$. Proceeding as for H_Q^+, it is straightforward to show that H_Q^- is given by (4.29). In the limiting case of $\alpha_3 = 0$, the poles $f_{\pm}$ both lie *on* the unit circle. The contour integral is evaluated by making vanishingly small indentations around the two poles. The same result is obtained by taking the case $\alpha_3 = 0$ in (4.28).

Appendix E

We require the integrals

$$I(z,\gamma,\xi_1) = \frac{1}{\mu}\int d\xi_2 \frac{1}{|\xi|^2}\frac{e^{-i\xi_2 z}}{\xi_2+\gamma\xi_1}$$

and

$$J(z,\gamma,\xi_1) = \frac{1}{2\mu(1-\nu)}\int d\xi_2 \frac{1}{|\xi|^4}\frac{e^{-i\xi_2 z}}{\xi_2+\gamma\xi_1},$$

interpreted as Cauchy principal values, when $z > 0$ and $\xi_1 > 0$, γ being real. The integrals are evaluated by considering ξ_2 as a complex variable. The contour is closed in the lower half-plane in the usual way, and in addition a semicircular indentation is incorporated into the contour to exclude the simple pole at $\xi_2 = -\gamma\xi_1$ on the real axis. The integral around the resulting contour may be evaluated in two ways. On deforming the contour to infinity, only the real axis and the indentation contribute to the contour integral. The contribution from the indentation is, on making its radius vanishingly small, just half the contribution that would be obtained were the associated pole entirely enclosed by a similarly oriented contour. The remaining integral along the real axis is then the required integral I or J, interpreted as a Cauchy principal value. Alternatively, the contour integral may be evaluated using Cauchy's residue theorem; the contour encloses a pole at $\xi_2 = -i\xi_1$, which is a single pole in the case of I and a double pole in the case of J. In either case the residue is easily evaluated. Equating in each case the results of the two methods of integration yields the formulae given in Section 6.2, following (6.15).

Appendix F

Here the invariance property of the dilatation, (6.28), is derived directly from (6.5). For a mismatch strain of the form given in (6.22), (6.5) yields:

$$\sigma^s_{kk} = -c_{kkr\beta}c_{i\alpha ll}f_m \int_R dx' G_{ir,\alpha\beta}(\mathbf{x}-\mathbf{x}') - c_{kkll}f_m\chi_R(\mathbf{x}).$$

For the case of isotropy of the tensor of elastic moduli, this reduces to

$$\sigma^s_{kk} = -4\mu^2\left(\frac{1+\nu}{1-2\nu}\right)^2 f_m \int_R dx' G_{\alpha\beta,\alpha\beta}(\mathbf{x}-\mathbf{x}') - 6\mu\left(\frac{1+\nu}{1-2\nu}\right) f_m\chi_R. \tag{F.1}$$

Now,

$$\begin{aligned}
\int_R dx' G_{\alpha\beta,\alpha\beta}(\mathbf{x}-\mathbf{x}') &= -\int_R dx' \frac{1}{4\pi^2}\iint d\xi_1 d\xi_2 \xi_\alpha\xi_\beta \hat{G}_{\alpha\beta}(\boldsymbol{\xi}) e^{-i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}')} \\
&= -\frac{1}{2\mu}\left(\frac{1-2\nu}{1-\nu}\right)\int_R dx'\delta(\mathbf{x}-\mathbf{x}') \\
&= -\frac{1}{2\mu}\left(\frac{1-2\nu}{1-\nu}\right)\chi_R(\mathbf{x}).
\end{aligned}$$

Here the first equality is obtained by performing and inverting a two-dimensional Fourier transform, and the second equality is obtained by using the explicit formula for $\hat{G}_{\alpha\beta}(\boldsymbol{\xi})$, given in (6.13). Hence (F.1) becomes

$$\sigma^s_{kk} = -4\mu\left(\frac{1+\nu}{1-\nu}\right) f_m\chi_R(\mathbf{x}).$$

Application of (6.28) now yields (6.29) immediately.

Bibliography

[1] O. Albrektsen, H. P. Meier, D. H. Arent, and H. W. M. Salemink. *Appl. Phys. Lett.*, 62:2105, 1993.

[2] T. Arakawa, S. Tsukamoto, Y. Nagamune, M. Nishioka, J. H. Lee, and Y. Arakawa. *Jpn. J. Appl. Phys.*, 32:L1377, 1993.

[3] D. J. Bacon, R. Bullough, and J. R. Willis. *Phil. Mag.*, 22:31, 1970.

[4] D. J. Bacon and A. G. Crocker. *Phil. Mag.*, 12:195, 1965.

[5] D. J. Bacon and P. P. Groves. In J. A. Simmons, R. de Wit, and R. Bullough, editors, *Fundamental Aspects of Dislocation Theory*, page 35. Nat. Bur. Stand. (U. S.), 1970. Spec. Publ. 317, I.

[6] I. Balslev. *Phys. Rev.*, 143:636, 1966.

[7] G. E. Beltz and L. B. Freund. *Phys. Stat. Sol. (a)*, 180:303, 1993.

[8] G. L. Bir and G. E. Pikus. *Symmetry and Strain-Induced Effects in Semiconductors.* Wiley: New York, 1974.

[9] R. Bullough. *Theory of Imperfect Crystalline Solids: Trieste Lectures 1970.* International Atomic Energy Agency, Vienna, 1971.

[10] R. Bullough and A. J. E. Foreman. *Phil. Mag.*, 9:315, 1964.

[11] M. Chandrasekhar and F. H. Pollak. *Phys. Rev. B*, 15:2127, 1977.

[12] K. H. Chang, P. K. Bhattacharya, and R. Gibala. *J. Appl. Phys.*, 66:2993, 1989.

[13] M. Comninou and J. Dundurs. *J. Elasticity*, 5:203, 1975.

[14] R. de Wit. *Solid-State Physics*, 10:249, 1960.

[15] J. D. Eshelby. *Phil. Trans. Roy. Soc.*, A244:87, 1951.

[16] J. D. Eshelby. *Proc. R. Soc. London*, A 241:376, 1957.

[17] J. D. Eshelby, W. T. Read, and W. Shockley. *Acta Metall.*, 1:251, 1953.

[18] J. Feder, K. C. Russell, J. Lothe, and G. M. Pound. *Adv. Phys.*, 15:111, 1966.

[19] E. A. Fitzgerald. *J. Vac. Sci. Technol. B*, 7:782, 1988.

[20] E. A. Fitzgerald, D. G. Ast, P. D. Kirchner, G. D. Pettit, and J. M. Woodall. *J. Appl. Phys.*, 63:693, 1988.

[21] E. A. Fitzgerald, E. A. Watson, R. E. Proano, D. G. Ast, P. D. Kirchner, G. D. Pettit, and J. M. Woodall. *J. Appl. Phys.*, 65:2220, 1989.

[22] L. B. Freund. *J. Mech. Phys. Solids*, 38:657, 1990.

[23] L. B. Freund. *J. Appl. Phys.*, 68:2073, 1990.

[24] T. J. Gosling. *J. Appl. Phys.*, 74:5415, 1993.

[25] T. J. Gosling, R. Bullough, S. C. Jain, and J. R. Willis. *J. Appl. Phys.*, 73:8267, 1993.

[26] T. J. Gosling, S. C. Jain, J. R. Willis, A. Atkinson, and R. Bullough. *Phil. Mag. A*, 66:119, 1992.

[27] T. J. Gosling and J. R. Willis. *Phil. Mag. A*, 69:65, 1994.

[28] T. J. Gosling and J. R. Willis. *J. Mech. Phys. Solids*, 42:1199, 1994.

[29] P. P. Groves and D. J. Bacon. *Phil. Mag.*, 22:83, 1970.

[30] W. Hagen and H. Strunk. *Appl. Phys.*, 17:85, 1978.

[31] J. P. Hirth and J. Lothe. *Theory of Dislocations.* Wiley: New York, 1982.

[32] D. C. Houghton. *J. Appl. Phys.*, 70:2136, 1991.

[33] R. Hull and J. C. Bean. *J. Vac. Sci. Technol. A*, 7:2580, 1989.

[34] R. Hull and J. C. Bean. *J. Vac. Sci. Technol. A*, 7:2580, 1989.

[35] R. Hull, J. C. Bean, D. Bahnck, L. J. Peticolas Jr., K. T. Short, and F. C. Unterwald. *J. Appl. Phys.*, 70:2052, 1991.

[36] R. Hull, J. C. Bean, L. Peticolas, and D. Bahnck. *Appl. Phys. Lett.*, 59:964, 1991.

[37] S. C. Jain. *Germanium-Silicon Strained Layers and Heterostructures.* Academic Press: Boston, 1994.

[38] S. C. Jain, T. J. Gosling, J. R. Willis, D. H. J. Totterdell, and R. Bullough. *Phil. Mag. A*, 65:1151, 1992.

[39] S. C. Jain and W. Hayes. *Semicond. Sci. Technol.*, 6:547, 1991.

[40] U. Jain, S. C. Jain, A. H. Harker, and R. Bullough. *J. Appl. Phys.*, 1994. Submitted for publication.

[41] W. A. Jesser and J. H. van der Merwe. *J. Appl. Phys.*, 75:872, 1994.

[42] S. V. Kamat and J. P. Hirth. *J. Appl. Phys.*, 67:6844, 1990.

[43] C. J. Kiely, J. I. Chyi, A. Rockett, and Morkoç. *Phil. Mag. A*, 60:321, 1989.

[44] E. P. Kvam, D. M. Maher, and C. J. Humphreys. *J. Mater. Res.*, 5:1900, 1990.

[45] L. D. Laude, F. H. Pollak, and M. Cardona. *Phys. Rev. B.*, 3:2623, 1971.

[46] A. Lefebvre, C. Herbeaux, C. Bouiller, and J. Di Persio. *Phil. Mag. Lett.*, 63:23, 1991.

[47] F. K. LeGoues, P. M. Mooney, and J. O. Chu. *Appl. Phys. Lett.*, 62:140, 1993.

[48] J. C. M. Li. *Phil. Mag.*, 10:1097, 1964.

[49] F. Louchet, D. Cochet Muchy, and Y. Brechet. *Phil. Mag. A*, 57:327, 1988.

[50] J. W. Matthews and A. E. Blakeslee. *J. Cryst. Growth*, 27:118, 1974.

[51] R. D. Mindlin. *Physics*, 7:195, 1936.

[52] T. Mura. *Phil. Mag.*, 8:843, 1963.

[53] J. Narayan and S. Sharan. *Mater. Sci. Eng. B*, 10:261, 1991.

[54] P. J. Orders and B. F. Usher. *Appl. Phys. Lett.*, 50:980, 1987.

[55] S. S. Orlov and V. L. Indenbom. *Sov. Phys. Crystallogr.*, 14:675, 1970.

[56] T. P. Pearsell, editor. *Strained-Layer Superlattices: Physics*, volume 32 of *Semiconductors and Semimetals.* Academic Press, 1990.

[57] D. D. Perovic and D. C. Houghton. *Proc. Mat. Res. Soc.*, 263:391, 1992.

[58] F. H. Pollak. *Strained-Layer Superlattices: Physics*, volume 32 of *Semiconductors and Semimetals*, chapter 2. Academic Press, 1990.

[59] J. R. Rice and G. E. Beltz. *J. Mech. Phys. Solids*, 42:333, 1994.

[60] A. Rocher, F. W. O. Da Silva, and C. Raisin. *Revue Phys. Appl.*, 25:957, 1990.

[61] G. Simmons and H. Wang. *Single Crystal Elastic Constants and Calculated Aggregate Properties.* M.I.T. Press: Cambridge, Massachusetts, 1971.

[62] K. R. Sprung, K. Wilke, G. Hetmann, J. Varrio, and M. Pessa. *Apopl. Phys. Lett.*, 62:2711, 1993.

[63] A. N. Stroh. *Phil. Mag.*, 3:625, 1958.

[64] A. N. Stroh. *J. Math. Phys.*, 41:77, 1962.

[65] T. C. T. Ting and D. M. Barnett. *Int. J. Solids Struct.*, 30:313, 1993.

[66] C. G. Tuppen and C. J. Gibbings. *J. Appl. Phys.*, 68:1526, 1990.

[67] C. G. Tuppen, C. J. Gibbings, and M. Hockly. *J. Crystal Growth*, 94:392, 1989.

[68] N. Usami, T. Mine, S. Fukatsu, and Y. Shiraki. *Appl. Phys. Lett.*, 64:1126, 1994.

[69] J. H. van der Merwe. *Surf. Sci.*, 31:198, 1972.

[70] C. G. Van der Walle and R. M. Martin. *Phys. Rev. B*, 34:5621, 1986.

[71] V. Volterra. *Annls. Ec. Norm.*, 24:401, 1907.

[72] I. Vurgaftman, J. M. Hinckley, and J. Singh. *IEEE Journal of Quantum Electronics*, 30:75, 1994.

[73] K. Wada, A. Kozen, Y. Hasumi, and J. Temmyo. *Appl. Phys. Lett.*, 54:436, 1989.

[74] J. R. Willis, S. C. Jain, and R. Bullough. *Phil. Mag. A*, 62:115, 1990.

[75] J. R. Willis, S. C. Jain, and R. Bullough. *Phil. Mag. A*, 64:629, 1991.

[76] T. Yamauchi, Y. Arakawa, and J. N. Schulman. *Surface Science*, 267:291, 1992.

[77] T. Yamauchi, T. Takahashi, J. N. Schulman, and Y. Arakawa. *IEEE Journal of Quantum Electronics*, 29:2109, 1993.